

\*\*\*\*\*

```

RESULT      3
ID   R51553 standard; Protein: 396 AA.
AC   R51553;
DT   09-JUN-1995    (first entry)
DE   Prepro human CBMP2A .
KW   Prepro human CBMP2A; cartilage and endochondrial bone formation;
KW   allograft repair; osteoarthritis; non-union fracture repair;
KW   osteogenesis; periodontal, dental and craniofacial reconstruction.
OS   Homo sapiens.
PN   US53457-A.
PD   11-OCT-1994.
PF   08-APR-1988;    179406.
PR   08-APR-1988;    PS-179406.
PR   15-AUG-1988;    US-232630.
PR   23-FEB-1989;    US-315342.
PR   17-OCT-1989;    US-422613.
PR   17-OCT-1989;    US-422699.
PR   22-FEB-1990;    US-483913.
PR   20-AUG-1990;    US-569920.
PR   07-SEP-1990;    US-579865.
PR   18-OCT-1990;    US-599543.
PR   18-OCT-1990;    US-600024.
PA   04-DEC-1990;    PS-621849.
PA   04-DEC-1990;    US-621988.
PA   22-FEB-1991;    US-660162.
PA   20-DEC-1991;    US-810560.
PA   28-JAN-1992;    US-827052.
PA   21-FEB-1992;    US-841846.
PA   18-DEC-1992;    US-993387.
PI   Kuberasampath T, Oppermann H, Ozkaynak E, Pang RHL;
PI   Rueger DC;
DR   N-PSDB: Q72709.
PT   Implantable device for inducing osteogenesis - comprises porous
PT   matrix contg. non-glycosylated dimeric, disulphide linked
PT   osteogenic protein.
PS   Disclosure; Columns 85-90; 128pp; English.
SC   Q72709 encodes R51653 prepro human CBMP2A, fragments of the
CC   related protein CBMP3 (R51642) consisting of residues 93-109 and
CC   77-83 are unglycosylated osteogenic polypeptides. These polypeptides
CC   can be disulphide bonded to form a dimer, which forms an essential
CC   component of an osteogenic protein. This protein is dispersed
CC   in a biodegradable matrix which can be implanted into a mammalian
CC   bone marrow cavity, here it can induce local cartilage, bone and
CC   endochondrial bone formation; and it can also accelerate allograft
CC   repair. This implant has the advantage of inducing all stages of
CC   bone formation and of having a higher specific activity than other
CC   known biosynthetic materials. The implant can be used to repair
CC   non-union fractures and cartilage; treat osteoarthritis; and aid
CC   in periodontal, dental or craniofacial reconstruction.
SQ   Sequence 396 AA.;

Query Match          30.6%; Score 67; DB 12; Length 396;
Best Local Similarity 36.8%; Pred.No. 2.61e+01;
Matches              7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;


Db       211 pavmrvtagghanhgfvrve 229
QY       |::|::|:
         7 PQILLWDPSQGQGFVVD 25


RESULT      4
ID   R85762 standard; Protein: 396 AA.
AC   R85762;
DT   20-JUN-1996    (first entry)
DE   Human CBMP2A.
KW   Human; osteogenic protein; hOP-1; murine; mOP-1; TGF-beta superfamily;
KW   transforming growth factor-beta; dimer; antibody; epitope; hippocampus;
KW   purification; implantable osteogenic device; bone formation; craniofacial;
KW   anomaly; skeletal; dental; endochondral bone formation;
KW   non-union fracture; cartilage repair; osteoarthritis.

```





CC congenital, trauma induced or oncologic resection induced craniofacial  
CC defects, and also is useful in cosmetic plastic surgery. It may be  
CC used in the treatment of periodontal disease and in other tooth  
CC repair processes. It may also be useful in the treatment of  
CC osteoporosis, wound healing (e.g. burns, incisions and ulcers) and  
CC related tissue repair, and may increase neuronal survival and be useful  
CC in the transplantation and treatment of conditions exhibiting a decrease  
CC in neuronal survival. It may be combined with other agents beneficial to  
CC the bone and/or cartilage defect, wound or tissue in question, e.g. EGF,  
CC PDGF, TGF-alpha, TGF-beta and insulin-like growth factor  
CC Sequence 396 AA;

Query Match 30.6%; Score 67; DB 7; Length 396;

Best Local Similarity 36.8%; Pred. No. 2.61e+01;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwttagghanhgfve 229

QY 7 PQILLWDPGQKQGQGFVVD 25

#### RESULT 7

ID W44303 standard; Protein; 396 AA.

AC W44303;

DT 27-MAY-1998 (first entry)

DE Human osteogenic protein CBMP2A.

KW Human; osteogenic protein; subunit; endochondral bone formation;

KW dimeric; recombinant protein.

OS Homo sapiens.

PN US5714589-A.

PD 03-FEB-1998.

PE 23-MAY-1995; 447570.

PR 21-FEB-1992; US-841646.

PR 08-APR-1988; US-179406.

PR 15-AUG-1988; US-232630.

PR 23-FEB-1989; US-315342.

PR 17-OCT-1989; US-422613.

PR 17-OCT-1989; US-422699.

PR 22-FEB-1990; US-483913.

PR 20-AUG-1990; US-569920.

PR 07-SEP-1990; US-579865.

PR 18-OCT-1990; US-595543.

PR 18-OCT-1990; US-600024.

PR 21-NOV-1990; US-616374.

PR 04-DEC-1990; US-621849.

PR 04-DEC-1990; US-621988.

PR 22-FEB-1991; US-660162.

PR 20-DEC-1991; US-810560.

PR 28-JAN-1992; US-827052.

PR 01-NOV-1993; US-147023.

PR 23-MAY-1995; US-447570.

PA (STVC ) STRYKER CORP.

PI Kuberasesampath T, Oppermann H, Ozkaynak E, Pang RHL,

PI Rueger DC;

DR WPI; 98-158353/14.

DR N-PSDB; V15206.

PT Extraction of osteogenic protein from mixture - using antibodies

PT specific for novel polypeptide chains useful as subunit(s) of

PT dimeric osteogenic protein(s)

PS Disclosure; Column 87-90; 127pp; English.

CC The present sequence represents a human osteogenic protein, which is  
CC used in the present invention. The present invention describes methods  
CC for selectively extracting an osteogenic protein (OP) from a mixture.

CC The method comprises: (a) exposing the mixture to an antibody that

CC specifically binds OP, separating the resulting antibody-protein complex

CC from the mixture, and dissociating the complex. In the methods OP

CC comprises a pair of oxidised subunits that are disulphide-bonded to form

CC a dimer, and one of the subunits has an amino acid (aa) sequence

CC sufficiently homologous to residues 335-431 of a 431 aa protein

CC designated OPs, sequence given in the specification. In dimeric form OP

CC is capable of inducing cartilage and endochondral bone formation in a

CC mammal when disposed within a matrix implanted in the mammal. The

CC methods are used for recovering the recombinant proteins from cell

CC cultures. 396 AA;  
SQ Sequence

Query Match 30.6%; Score 67; DB 29; Length 396;

Best Local Similarity 36.8%; Pred. No. 2.61e+01;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwttagghanhgfve 229

QY 7 PQILLWDPGQKQGQGFVVD 25

#### RESULT 8

ID W24849 standard; protein; 396 AA.

AC W24849;

DT 22-JAN-1998 (first entry)

DE Human bone morphogenic protein (BMP) 2A.

KW bone morphogenic protein; BMP 2A; cartilage; periodontal disease;

KW tissue repair; osteoporosis; treatment; oligonucleotide probe.

OS Homo sapiens.

FS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..23

FT Peptide /note= "pre-portion of BMP-2A"

FT Peptide 24..282

FT Protein /note= "pro-portion of BMP-2A"

FT Protein 283..396

FT Protein /note= "mature portion of BMP-2A"

FT Protein 299..396

FT Protein /note= "BMP-2A contains at least this portion of

FT Peptide 384..392

FT Peptide /note= "tryptic peptide fragment"

PN US5631142-A.

PD 20-MAY-1997.

PF 07-SEP-1993; 118363.

PR 11-JUL-1989; US-378537.

PR 17-DEC-1986; US-943332.

PR 20-MAR-1987; US-028285.

PR 08-APR-1988; US-179100.

PR 18-MAY-1992; US-884353.

PR 07-SEP-1993; US-118363.

PR 01-JUL-1986; US-880776.

(GENY ) GENETICS INST INC.

PI Rosen VA, Wang EA, Wozney JM;

DR WPI; 97-288573/26.

DR N-PSDB; T78941.

PT Production of human bone morphogenic protein 2A or 2B in cell

PT culture - useful inducing bone or cartilage production, in wound

PT healing and tissue repair

PS Claim 2; Fig 2; 22pp; English.

CC This is human bone morphogenic protein (BMP) 2A. Oligonucleotide probes  
CC were synthesised based on a partially sequenced BMP-2A protein isolated  
CC from ground bovine powder. The probes were used to screen a bovine liver  
CC DNA library to obtain the encoding DNA sequence. The DNA was used to  
CC screen a U-2 OS human cell line cDNA library to obtain human BMP-2A cDNA.  
CC BMPs can be used to induce bone and cartilage formation, and in wound  
CC healing and tissue repair. They can also be used for treating periodontal  
CC disease or osteoporosis.

SQ Sequence 396 AA;

Query Match 30.6%; Score 67; DB 25; Length 396;

Best Local Similarity 36.8%; Pred. No. 2.61e+01;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwttagghanhgfve 229

QY 7 PQILLWDPGQKQGQGFVVD 25

#### RESULT 9

ID R44747 standard; Protein; 396 AA.

AC R44747;

DT 06-JUN-1994 (first entry)



```

Db 211 pavmrwtgqghanhgfve 229
QY 7 POILLWDPGQKGGFVVD 25

RESULT 12
ID R14241 standard; Protein: 396 AA.
AC R14241;
DT 13-JAN-1992 (first entry)
DE Human BMP-2A encoded by lambda U20S-39.
KW Bone; cartilage; osteoinductive protein.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 384..392
   /label= tryptic
PN US013649-A.
PD 07-MAY-1991.
PF 08-APR-1988; 179100.
PR 01-JUL-1986; US-880776.
PR 04-APR-1988; US-179100.
PA (GENE-) Genetics Inst Inc.
DR WPI: 91-309401/42.
DR N-PSDB; Q14036.
PT New DNA sequences encoding osteo-inductive protein - useful for
PT stimulating bone and cartilage re formation e.g. for wound healing
PT and tissue repair.
PS Disclosure; Fig 2; 20pp; English.
CC The sequence was deduced from clone lambda U20S-39 (ATCC 40345)
CC which was isolated from a cDNA library prep. using human cell line
CC U-2 OS RNA. It is one of two classes of clones isolated
CC distinguished by hybridisation characteristics. This class,
CC designated hBMP-2A (previously BMP-2 and BMP-2 Class I) hybridised
CC strongly with the probe which was derived from the bovine BMP-2A
CC sequence, and showed strong homology with that sequence. The second
CC class, hBMP-2B (previously BMP-4 and BMP-2 Class II) hybridised only
CC weakly and was less homologous, esp. at the 3' end. The gene prod.
CC is an osteoinductive protein useful for inducing bone/cartilage
CC repair, wound healing and tissue repair. Typical applications
CC include healing of bone fractures; improved fixation of artificial
CC joints; in cosmetic plastic surgery; and in treatment of periodontal
CC disease, burns, incisions, ulcers etc.
CC See also R14240 and R14242.
SQ Sequence 396 AA;

Query Match 30.6%; Score 67; DB 3; Length 396;
Best Local Similarity 36.8%; Pred. No. 2.61e+01;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwtgqghanhgfve 229
QY 7 POILLWDPGQKGGFVVD 25

RESULT 13
ID R47255 standard; Protein: 396 AA.
AC R47255;
DT 15-AUG-1994 (first entry)
DE Pre-pro-BMP2.
KW Human; hippocampus; osteogenic protein; OP-1; subunit; dimer;
KW morphogenic activity; cysteine; morphogen; family; pro-region;
KW complex; soluble; aqueous solvent; therapeutic composition;
KW symptom-alleviating; co-factor; antibody; diagnosis; assay;
KW quantitate; mature.
OS Homo sapiens.
PN W09403600-A.
PD 17-FEB-1994.
PF 29-JUL-1993; U07189.
PR 31-JUL-1992; US-923780.
PR 04-MAR-1993; US-029335.
PR 31-MAR-1993; US-040510.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PI Jones WK, Kuberasingh T, Oppermann H, Ozkaynak E;

```

```

PI Rueger DC, Tucker RF;
DR WPI: 94-065689/08.
PT Morphogenic protein soluble complex - for regeneration of tissue
PT in mammals and diagnosing tissue disorders
PS Claim 3; Page 81-83; 120pp; English.
CC The sequences given in R47254-64 represent osteogenic proteins. The
CC mature osteogenic proteins may be used as at least one subunit in the
CC dimeric protein of the invention. This dimeric protein comprises a
CC pair of protein subunits which are associated to give a structure
CC with morphogenic activity. Each subunit comprises more than 100
CC amino acids having a pattern of cysteine residues characteristic of
CC the morphogen family. Each subunit comprises a mature form of a
CC subunit of a member of the morphogen family, non-covalently complexed
CC with a peptide comprising a pro-region of a morphogenic family member,
CC to form a complex more soluble in aqueous solvents than the uncomplexed
CC subunits. The dimeric protein is useful in a therapeutic composition,
CC pref. also containing a symptom-alleviating co-factor. The protein
CC and corresponding antibody may be used in diagnostic assays, eg. to
CC quantitate the amount of mature and soluble forms of morphogenic
CC proteins produced.
SQ Sequence 396 AA;

Query Match 30.6%; Score 67; DB 9; Length 396;
Best Local Similarity 36.8%; Pred. No. 2.61e+01;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwtgqghanhgfve 229
QY 7 POILLWDPGQKGGFVVD 25

RESULT 14
ID R15474 standard; protein: 400 AA.
AC R15474;
DT 12-MAR-1992 (first entry)
DE BMP2A/2B fusion protein.
KW Expression; precursor; bone morphogenic protein; osteogenic.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..267
   /label= BMP-2A_precursor
FT region 268..400
   /label= mature_BMP-2B
FT- W09118047-A.
PN 28-NOV-1991.
PD 20-MAY-1991; U03540.
PR 24-MAY-1990; US-528300.
PA (GETH ) GENENTECH INC.
PI Hammonds Rg, Mason AJ;
DR WPI: 91-369212/50.
PT DNA encoding mature BMP-2, contg. upstream non-BMP-2 precursor -
PT giving improved expression and higher prodn. of mature BMP-2
PT protein than using native BMP-2 precursor
PS Disclosure; Fig 2; 35pp; English.
CC Improved expression levels of BMP-2 DNA in mammalian cells, over
CC that attainable using the BMP-2 precursor protein portion that is
CC native to the BMP are achieved without amplifying the DNA. This
CC improved expression and higher prodn. of BMP protein results in
CC sufficient quantities to test for its osteogenic activity, and also
CC to produce it on a commercially viable scale.
CC See also R15474, Q1524-37, Q15286 and Q15350..
SQ Sequence 400 AA;

Query Match 30.6%; Score 67; DB 3; Length 400;
Best Local Similarity 36.8%; Pred. No. 2.61e+01;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 pavmrwtgqghanhgfve 229
QY 7 POILLWDPGQKGGFVVD 25

RESULT 15

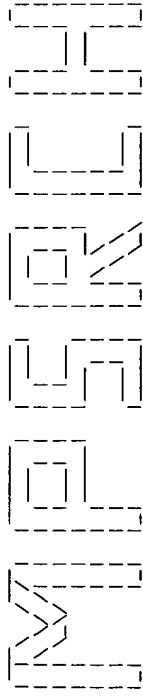
```

Db 211 pavmrwtagghanhgfvve 229  
| : | : | : |||:  
QY 7 PQILLWDPSGQKQGFFVD 25

Search completed: Mon Aug 23 13:12:13 1999  
Job time : 29 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:11:11 1999; MasPar time 8.97 Seconds  
Tabular output not generated. 129.496 Million cell updates/sec

Title: >US-09-049-696-45  
Description: (1-29) from US09049696.pap  
Perfect Score: 219  
Sequence: 1 TWTTPQILLWDPGKQGGFVVDKNTK 29

Scoring table: PAM 150  
Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 30.276; Variance 47.289; scale 0.640

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	148	67.6	913	3	gob-5 protein - Mouse	1.18e-16
2	78	35.6	794	2	hypothetical protein	3.33e-02
3	78	35.6	820	2	Lu-ECAM-1 protein - b	3.33e-02
4	78	35.6	905	2	Lu-ECAM-1 protein - b	3.33e-02
5	75	34.2	144	2	conserved hypothetical	1.13e-01
6	73	33.3	715	1	77K alpha trans-induc	2.51e-01
7	71	32.4	353	2	bone morphogenetic pr	5.52e-01
8	70	32.0	393	2	bone morphogenetic pr	8.13e-01
9	69	31.5	351	2	bone morphogenetic pr	1.19e+00
10	69	31.5	394	2	bone morphogenetic pr	1.19e+00
11	68	31.1	217	2	uracil-DNA glycosylas	1.75e+00
12	68	31.1	372	2	probable membrane pro	1.75e+00
13	67	30.6	396	1	bone morphogenetic pr	2.55e+00
14	66	30.1	138	2	xpsi protein - Xantho	3.71e+00
15	66	30.1	357	2	naringenin 3-dioxygen	3.71e+00
16	66	30.1	357	2	naringenin 3-dioxygen	3.71e+00
17	66	30.1	475	2	metalloproteinase G (	3.71e+00
18	66	30.1	1052	2	protein-tyrosine kina	3.71e+00
19	66	30.1	1068	2	focal adhesion kinase	3.71e+00
20	65	29.7	229	1	uracil-DNA glycosylas	5.37e+00
21	65	29.7	642	2	hypothetical protein	5.37e+00
22	65	29.7	1052	2	focal adhesion kinase	5.37e+00
23	64	29.2	263	2	hypothetical protein	7.74e-00

24	64	29.2	357	2	S38338	naringenin 3-dioxygen	7.74e+00
25	64	29.2	359	2	S16780	naringenin 3-dioxygen	7.74e+00
26	64	29.2	359	2	A42110	flavanone 3 beta-hydr	7.74e+00
27	64	29.2	374	2	S57750	naringenin 3-dioxygen	7.74e+00
28	64	29.2	382	2	S1281	cytochrome c-type bio	7.74e+00
29	64	29.2	577	2	S77715	cytochrome c-type bio	7.74e+00
30	64	29.2	589	2	S38799	hypothetical protein	7.74e+00
31	64	29.2	636	2	S76730	hypothetical protein	7.74e+00
32	64	29.2	710	2	S54152	sepB protein - EmERIC	7.74e+00
33	64	29.2	846	1	PNECA	penicillin amidase (E	7.74e+00
34	63	28.8	434	2	S73331	hypothetical protein	1.11e+01
35	63	28.8	475	2	A74807	opcA protein - Synec	1.11e+01
36	63	28.8	686	2	S38235	microtubule-associate	1.11e+01
37	63	28.8	1005	2	I40588	transposase (Tn5401)	1.11e+01
38	62	28.3	372	2	T03385	naringenin 3-dioxygen	1.59e+01
39	62	28.3	443	2	E64725	yadU protein - Escher	1.59e+01
40	61	27.9	819	1	TNBE11	91.8K alpha trans-ind	2.27e+01
41	61	27.9	898	2	S65474	pyruvate dehydrogenas	2.27e+01
42	61	27.9	938	2	A39160	transcription activat	2.27e+01
43	61	27.9	1010	2	I40329	brkA prtoein - Bordet	2.27e+01
44	61	27.9	1184	2	A55184	fibulin-2 precursor	2.27e+01
45	60	27.4	478	2	S04784	glucose dehydrogenase	3.21e+01

ALIGNMENTS

RESULT 1  
ENTRY JG0168 #type complete  
TITLE gob-5 protein - Mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 18-Mar-1999 #sequence\_revision 18-Mar-1999 #text\_change 18-Mar-1999  
ACCESSIONS JG0168  
REFERENCE JG0168  
#authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.  
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351  
#title Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice.  
#accession JG0168  
##status preliminary  
##residues 1-913 ##label KOM  
##cross-references DDBJ:AB016592  
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 67.6%; Score 148; DB 3; Length 913;  
Best Local Similarity 62.1%; Pred. No. 1.18e-16;  
Matches 18; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 522 TWTTHPTFIWDPGVEONGFILDITTK 550  
QY 1 TWTTPQILLWDPGKQGGFVVDKNTK 29

RESULT 2  
ENTRY T02171 #type complete  
TITLE hypothetical protein Lu-ECAM-1 - bovine  
ORGANISM #formal\_name Bos primigenius taurus #common\_name cattle  
DATE 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 05-Mar-1999  
ACCESSIONS T02171  
REFERENCE Z14590

#authors Eble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.; Levine, R.; Goodwin, A.; Pauli, B.U.  
#submission submitted to the EMBL Data Library, April 1997  
#description Cloning and characterization of Lu-ECAM-1 suggest it is an endothelial chloride channel.

#accession T02171  
##status preliminary; translated from GB/EMBL/DDBJ  
##molecule\_type mRNA  
##residues 1-794 ##label ELB  
##cross-references EMBL:AF001262; NID:g2623764; PID:g2623765  
##experimental\_source lung  
SUMMARY #length 794 #molecular-weight 88509 #checksum 5853

```

Query Match      35.6%; Score 78; DB 2; Length 794;
Best Local Similarity 61.1%; Pred. No. 3.33e-02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIOKPEIVLDPKPKKK 540
|||||:|:|:|:|:|
Qy 1 TWTTPPQILLWDPGQK 18

RESULT 3
ENTRY #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02152
REFERENCE T02152
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Chany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.U.
#description submitted to the EMBL Data Library, April 1997
#accession Cloning and characterization of Lu-ECAM-1 suggest it is an
#status endotheial chloride channel.
#molecule_type T02152 preliminary; translated from GB/EMBL/DDBJ
#residues 1-820 #label ELB
#cross-references EMBL:AF001263; NID:g2623766; PID:g2623767
#experimental_source lung
SUMMARY #length 820 #molecular-weight 91464 #checksum 3104

Query Match      35.6%; Score 78; DB 2; Length 820;
Best Local Similarity 61.1%; Pred. No. 3.33e-02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIOKPEIVLDPKPKKK 540
|||||:|:|:|:|:|
Qy 1 TWTTPPQILLWDPGQK 18

RESULT 4
ENTRY #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02205
REFERENCE T02205
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Chany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.U.
#description submitted to the EMBL Data Library, April 1997
#accession Cloning and characterization of Lu-ECAM-1 suggest it is an
#status endotheial chloride channel.
#molecule_type T02205 preliminary; translated from GB/EMBL/DDBJ
#residues 1-905 #label ELB
#cross-references EMBL:AF001261; NID:g2623762; PID:g2623763
#experimental_source lung
SUMMARY #length 905 #molecular-weight 101005 #checksum 5557

Query Match      35.6%; Score 78; DB 2; Length 905;
Best Local Similarity 61.1%; Pred. No. 3.33e-02;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIOKPEIVLDPKPKKK 540
|||||:|:|:|:|:|
Qy 1 TWTTPPQILLWDPGQK 18

RESULT 5
ENTRY #type complete
TITLE conserved hypothetical protein AF2387 - Archaeoglobus

```

```

fulgidus
#formal_name Archaeoglobus fulgidus
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
ACCESSIONS C69548
REFERENCE A69250
#authors Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.;
Peterson, J.D.; Richardson, D.L.; Kierlavage, A.R.; Graham,
D.E.; Kyripides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.;
Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
T.; Artach, P.; Raine, B.P.; Sykes, S.M.; Sadow, P.W.;
D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.;
Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
C.R.; Venter, J.C.
#journal Nature (1997) 390:364-370
#title The complete genome sequence of the hyperthermophilic,
sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references MUID:98049343
#accession C69548
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-144 #label KLE
#cross-references GB:AE001111; GB:AE000782; NID:g2689434; PID:g2650702;
TIGR:AF2387
SUMMARY #length 144 #molecular-weight 16212 #checksum 5934

Query Match      34.2%; Score 75; DB 2; Length 144;
Best Local Similarity 46.7%; Pred. No. 1.13e-01;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 49 ILLWASRRRRRGVF 63
||||:|:|:|:|:|
Qy 9 ILLWDPGQKRGVF 23

RESULT 6
ENTRY #type complete
TITLE 77K alpha trans-inducing protein - human herpesvirus 1
ORGANISM #formal_name human herpesvirus 1
DATE 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change
ACCESSIONS B26133
REFERENCE A93031
#authors McKnight, J.L.C.; Pellett, P.E.; Jenkins, F.J.; Roizman, B.
#journal J. Virol. (1987) 61:992-1001
#title Characterization and nucleotide sequence of two herpes
simplex virus 1 genes whose products modulate
alpha-trans-inducing factor-dependent activation of alpha
genes.
#cross-references MUID:87141362
#accession B26133
#molecule_type DNA
#residues 1-715 #label MCK
#cross-references GB:M15621; NID:g330056; PID:g330058
COMMENT This protein may enhance transcriptional activity of the alpha
genes.
CLASSIFICATION #superfamily herpesvirus 77K alpha trans-inducing protein
KEYWORDS trans-inducing protein; transcription regulation
SUMMARY #length 715 #molecular-weight 77357 #checksum 1578

Query Match      33.3%; Score 73; DB 1; Length 715;
Best Local Similarity 37.0%; Pred. No. 2.51e-01;
Matches 10; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Db 431 TWSAGPPDDMARGPGGHRAGGTGK 457
||::|:|:|:|:|:|

```



```
QY 1 TWTT-QPQILLWDPGKQGGFVVDK 26

RESULT 7
ENTRY bone morphogenetic protein 2 - chicken (fragment)
TITLE #formal_name Gallus gallus #common_name chicken
ORGANISM 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
DATE 23-Feb-1997

ACCESSIONS I50607
REFERENCE I50607
#authors Francis, P.H.; Richardson, M.K.; Brickell, P.M.; Tickle, C.
#journal Development (1994) 120:209-218
#title Bone morphogenetic proteins and a signalling pathway that
controls patterning in the developing chick limb.
#cross-references MUID:94163974
#accession I50607
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-353 #label FRA
#cross-references EMBL:X75914; NID:g472927; PID:g472928
GENETICS
#gene Bmp-2
CLASSIFICATION #superfamily inhibin
SUMMARY #length 353 #checksum 7551

Query Match 32.48; Score 71; DB 2; Length 353;
Best Local Similarity 47.4; Pred. No. 5.52e-01;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Db 168 PAVLRWIAHQGNHGFVVE 186
| : | : | : | : |
QY 7 PQILLWDPGKQGGFVVD 25

RESULT 8
ENTRY S37073 #type complete
TITLE bone morphogenetic protein 2 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
08-Sep-1997

ACCESSIONS S37073
REFERENCE S37073
#authors Feng, J.Q.; Chen, D.; Feng, M.; Harris, M.A.; Mundy, G.R.;
Harris, S.E.
#submission submitted to the EMBL Data Library, September 1993
#description cDNA sequence of fetal rat calvarial osteoblast bone
morphogenetic protein 2.
#accession S37073
#status preliminary
#molecule_type mRNA
#residues 1-393 #label FEN
#cross-references EMBL:Z25868; NID:g397950; PID:g397951
CLASSIFICATION #superfamily inhibin
SUMMARY #length 393 #molecular-weight 44382 #checksum 7639

Query Match 32.08; Score 70; DB 2; Length 393;
Best Local Similarity 36.8; Pred. No. 8.13e-01;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 208 PAVMRWTAQGHNHGFVVE 226
| : | : | : |
QY 7 PQILLWDPGKQGGFVVD 25

RESULT 9
ENTRY A34201 #type complete
TITLE bone morphogenetic protein homolog 2A - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 01-Jun-1990 #sequence_revision 01-Jun-1990 #text_change
15-Jun-1996

ACCESSIONS A34201
REFERENCE A34201
#authors Crouzet, J.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez,
M.C.; Blanche, F.; Thibaut, D.; Debussche, L.
#journal J. Bacteriol. (1990) 172:5980-5990
#title Genetic and sequence analysis of an 8.7-kilobase Pseudomonas
denitrificans fragment carrying eight genes involved in
transformation of precorrin-2 to cobyrinic acid.
#cross-references MUID:91008976
```

```

#authors Dickinson, M.E.; Kobrin, M.S.; Silan, C.M.; Kingsley, D.M.;
Justice, M.J.; Miller, D.A.; Ceci, J.D.; Lock, L.F.; Lee,
A.; Buchberg, A.M.; Siracusa, L.D.; Lyons, K.M.; Derynck,
R.; Hogan, B.L.M.; Copeland, N.G.; Jenkins, N.A.
#journal Genomics (1990) 6:505-520
#title Chromosomal localization of seven members of the murine
TGF-beta superfamily suggests close linkage to several
morphogenetic mutant loci.
#cross-references MUID:90228966
#accession A34201
#status preliminary; nucleic acid sequence not shown; not
compared with conceptual translation
#molecule_type mRNA
#residues 1-351 #label DIC
CLASSIFICATION #superfamily inhibin
SUMMARY #length 351 #molecular-weight 39863 #checksum 9879

Query Match 31.5; Score 69; DB 2; Length 351;
Best Local Similarity 36.8; Pred. No. 1.19e+00;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Db 209 PAVMRWTTQGHNHGFVVE 227
| : | : | : |
QY 7 PQILLWDPGKQGGFVVD 25

RESULT 10
ENTRY S45355 #type complete
TITLE bone morphogenetic protein-2 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 10-Dec-1994 #sequence_revision 17-Nov-1995 #text_change
03-May-1996

ACCESSIONS S45355
REFERENCE S45355
#authors Feng, J.Q.; Harris, M.A.; Ghosh-Choudhury, N.; Feng, M.;
Mundy, G.R.; Harris, S.E.
#journal Biochim. Biophys. Acta (1994) 1218:221-224
#title Structure and sequence of mouse bone morphogenetic protein-2
gene (BMP-2): comparison of the structures and promoter
regions of BMP-2 and BMP-4 genes.
#cross-references MUID:94289485
#accession S45355
#status preliminary
#molecule_type DNA
#residues 1-394 #label FEN
CLASSIFICATION #superfamily inhibin
SUMMARY #length 394 #molecular-weight 44509 #checksum 1098

Query Match 31.5; Score 69; DB 2; Length 394;
Best Local Similarity 36.8; Pred. No. 1.19e+00;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Db 209 PAVMRWTTQGHNHGFVVE 227
| : | : | : |
QY 7 PQILLWDPGKQGGFVVD 25

RESULT 11
ENTRY I36145 #type fragment
TITLE uracil-DNA glycosylase (EC 3.2.2.-) - Pseudomonas sp.
ORGANISM (fragment)
DATE 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change
25-May-1996

ACCESSIONS I36145
REFERENCE I36145
#authors Crouzet, J.; Cameron, B.; Cauchois, L.; Rigault, S.; Rouyez,
M.C.; Blanche, F.; Thibaut, D.; Debussche, L.
#journal J. Bacteriol. (1990) 172:5980-5990
#title Genetic and sequence analysis of an 8.7-kilobase Pseudomonas
denitrificans fragment carrying eight genes involved in
transformation of precorrin-2 to cobyrinic acid.
#cross-references MUID:91008976
```

```

#note      the source is designated as Pseudomonas denitrificans
#accession 136145
#status    preliminary
#molecule_type DNA
#residues  1-217 #label CRO
##cross-references GB:W32224
##note     this gene was not noted by the authors but was
           translated after correction for a likely frameshift as
           suggested by Bobby Baum; the location of the
           frameshift was assigned within a range of three codons
           on the basis of a multiple sequence alignment
CLASSIFICATION #superfamily uracil-DNA glycosylase
KEYWORDS       DNA repair; glycosidase; hydrolase
SUMMARY        #length 217 #checksum 294

Query Match      31.1%; Score 68; DB 2; Length 217;
Best Local Similarity 36.8%; Pred. No. 1.75e+00;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 143 HPVVFMLWGSYAKKAAFY 161
QY 5 QPQILLWDPGQKGGFV 23

RESULT 12
ENTRY   B64819      #type complete
TITLE   Probable membrane protein yb1R - Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE    12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
        13-Sep-1998
ACCESSIONS B64819
REFERENCE  A64720
#authors   Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
           Burdand, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
           Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
           Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
           Y.
#journal   Science (1997) 277:1453-1462
#title     The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession B64819
#status    nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues  1-372 #label BLAT
##cross-references GB:AE000184; GB:U000096; NID:gl787036; PID:gl787039;
           UWGP:b0818
#experimental_source strain K-12, substrain MG1655

GENETICS
#gene      yb1R
KEYWORDS   Transmembrane protein
FEATURE    17-33      #domain transmembrane #status predicted #label TM1\
           81-97      #domain transmembrane #status predicted #label TM2\
           100-116    #domain transmembrane #status predicted #label TM3\
           166-182    #domain transmembrane #status predicted #label TM4\
           202-218    #domain transmembrane #status predicted #label TM5\
           221-237    #domain transmembrane #status predicted #label TM6\
           241-257    #domain transmembrane #status predicted #label TM7\
           301-317    #domain transmembrane #status predicted #label TM8\
           352-368    #domain transmembrane #status predicted #label TM9\
SUMMARY    #length 372 #molecular-weight 41168 #checksum 2942

Query Match      31.1%; Score 68; DB 2; Length 372;
Best Local Similarity 38.1%; Pred. No. 1.75e+00;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Db 143 PQNLIWGRSGLSFAGFIAQM 163
QY 6 PQQILLWDPGQKGGFVVDK 26

RESULT 13
ENTRY   BMH2
#type complete

```

```

TITLE      bone morphogenetic protein 2 precursor - human
ALTERNATE_NAMES
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change
           05-Sep-1997
ACCESSIONS B37278; PC2178
REFERENCE  A37278
#authors   Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.;
           Whitters, M.J.; Kriz, R.W.; Hewick, R.M.; Wang, E.A.
#journal   Science (1988) 242:1528-1534
#title     Novel regulators of bone formation: molecular clones and
           activities.
#cross-references MUID:89072730
#accession B37278
#molecule_type mRNA
#residues  1-396 #label WO2
##cross-references GB:M22489; NID:gl79501; PID:gl79502
REFERENCE  PC2178
#authors   Ishida, N.; Tsujimoto, M.; Kanaya, T.; Shimamura, A.;
           Tsuruoka, N.; Kodama, S.; Katayama, T.; Oikawa, S.; Matsui,
           M.; Nakanishi, T.; Kobayashi, J.; Nakazato, H.
#journal   J. Biochem. (1994) 115:279-285
#title     Expression and characterization of human bone morphogenetic
           protein-2 in silkworm larvae infected with recombinant
           Bombyx mori nuclear polyhedrosis virus.
#accession PC2178
#molecule_type protein
#residues  290-295,'X'/297-304 #label ISH
#experimental_source cell line BOMO-15A1IC
REFERENCE  A56729
#authors   Rathore, S.; Hammerstone, K.M.; Dansereau, S.; Porter, T.J.
#journal   Protein Sci. (1995) 4[Suppl.2]:443S
#title     N-terminal isoforms of recombinant human bone morphogenetic
           protein (rhBMP-2) are active in vitro and in vivo.
#contents  annotation of amino ends of mature forms; dimers with long
           determination of form chains have reduced activity
COMMENT    This hormone is capable of inducing bone formation at ectopic
           morphological locations.
GENETICS
#gene      GDB:BMP2; BMP2A
#cross-references GDB:125204; OMIM:112261
#map_position 20p12-20p12
COMPLEX    homodimer, disulfide linked
CLASSIFICATION
KEYWORDS   bone; dimer; glycoprotein; pyroglutamic acid
FEATURE    1-23      #domain signal sequence #status predicted #label SIG\
           24-265    #domain propeptide #status predicted #label PRO\
           266-396  #product bone morphogenetic protein 2, long form #status
           predicted #label MATL\
           283-396  #product bone morphogenetic protein 2 #status predicted
           #label MAT\
           135,163,164,200 #binding_site carbohydrate (Asn) (covalent) #status
           predicted\
           283      #modified_site pyrrolidone carboxylic acid (Gln) (in
           mature form) #status experimental\
           338      #binding_site carbohydrate (Asn) (covalent) #status
           experimental
SUMMARY    #length 396 #molecular-weight 44702 #checksum 9850

Query Match      30.6%; Score 67; DB 1; Length 396;
Best Local Similarity 36.8%; Pred. No. 2.55e+00;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 PAYMRTAQGHANHGFFVE 229
QY 7 PQQILLWDPGQKGGFVVD 25

RESULT 14
ENTRY   SL7941
#type complete
TITLE    xpsi protein - Xanthomonas campestris pv. campestris

```

ORGANISM #formal\_name Xanthomonas campestris pv. campestris  
DATE 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change  
09-Sep-1997

ACCESSIONS S17941  
REFERENCE S17937  
#authors Dums, F.; Dow, J.M.; Daniels, M.J.  
#journal Mol. Gen. Genet. (1991) 229:357-364  
#title Structural characterization of protein secretion genes of the  
bacterial phytopathogen Xanthomonas campestris pathovar  
campestris: relatedness to secretion systems of other  
gram-negative bacteria.

#cross-references MUID:92049233  
#accession S17941  
#molecule\_type DNA  
#residues 1-138 #label DUM  
#cross-references EMBL:X59079; NID:g48896; PID:g48901

GENETICS  
#gene xpsi  
SUMMARY #length 138 #molecular-weight 15125 #checksum 7522

Query Match 30.1%; Score 66; DB 2; Length 138;  
Best Local Similarity 63.6%; Pred. No. 3.71e+00;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DB 113 WGEQPNORLLW 123  
| | | | |  
QY 2 WTQPPQILLW 12

RESULT 15  
ENTRY S61415 #type complete  
TITLE naringenin 3-dioxygenase (EC 1.14.11.9) - alfalfa  
ALTERNATE\_NAMES flavanone 3-hydroxylase  
ORGANISM #formal\_name Medicago sativa #common\_name alfalfa  
DATE 27-Apr-1996 #sequence\_revision 06-Sep-1996 #text\_change  
12-Feb-1999

ACCESSIONS S61415; S44244  
REFERENCE S61415  
#authors Chartier, B.; Coronado, C.; Kondorosi, A.; Ratet, P.  
#journal Plant Mol. Biol. (1995) 29:773-786  
#title Molecular characterization and expression of alfalfa  
(Medicago sativa L.) flavanone-3-hydroxylase and  
dihydroflavonol-4-reductase encoding genes.

#cross-references MUID:96128019  
#accession S61415  
#molecule\_type DNA  
#residues 1-357 #label CHA  
#cross-references EMBL:X78994; NID:g475958; PID:g475959

GENETICS  
#introns 121/3; 264/3  
CLASSIFICATION #superfamily 1-aminocyclopropane-1-carboxylate oxidase  
KEYWORDS oxidoreductase  
SUMMARY #length 357 #molecular-weight 40607 #checksum 8579

Query Match 30.1%; Score 66; DB 2; Length 357;  
Best Local Similarity 50.0%; Pred. No. 3.71e+00;  
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

DB 98 PEEKLRFDMGGKGGFVVSSH 119  
| | | | |  
QY 6 PPQILLWDFSGKGGFVVDKN 27

Search completed: Mon Aug 23 13:11:25 1999  
Job time : 14 secs.

**THIS PAGE BLANK (USPTO)**

\*\*\*\*\*

W P E R L F (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:10:04 1999; MasPar time 6.36 Seconds  
Tabular output not generated. 128.799 Million cell updates/sec

Title: >US-09-049-696-45  
Description: (1-29) from US09049696.pep  
Perfect Score: 219  
Sequence: 1 TWTTPPQILLMDPSGQGGFVVDKNTK 29  
Scoring table: PAM 150  
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 30.941; Variance 43.595; scale 0.710

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	78	35.6	903	1	ECLC_BOVIN	EPITHELIAL CHLORIDE CH	8.54e-03
2	73	33.3	715	1	AT12_HSV1F	ALPHA TRANS-INDUCING F	7.70e-02
3	71	32.4	353	1	BMP2_CHICK	BONE MORPHOGENETIC PRO	1.81e-01
4	70	32.0	393	1	BMP2_RAT	BONE MORPHOGENETIC PRO	2.76e-01
5	69	31.5	394	1	BMP2_MOUSE	BONE MORPHOGENETIC PRO	4.19e-01
6	68	31.1	217	1	UNG_PSEDE	URACIL-DNA GLYCOSYLASE	6.34e-01
7	68	31.1	372	1	YBIR_ECOLI	HYPOTHETICAL 41.2 KD P	6.34e-01
8	67	30.6	396	1	BMP2_HUMAN	BONE MORPHOGENETIC PRO	9.55e-01
9	66	30.1	138	1	GSPI_XANCP	GENERAL SECRETION PATH	1.43e+00
10	66	30.1	475	1	PTG_ERMCH	SECRETED PROTEASE G PR	1.43e+00
11	66	30.1	1052	1	FAK1_XENLA	FOCAL ADHESION KINASE	1.43e+00
12	65	29.7	228	1	UNG_ECOLI	URACIL-DNA GLYCOSYLASE	2.14e+00
13	65	29.7	1052	1	FAK1_HUMAN	FOCAL ADHESION KINASE	2.14e+00
14	65	29.7	1058	1	PMAL1_DICDI	PROBABLE PLASMA MEMBRA	2.14e+00
15	64	29.2	263	1	YMPF_ECOLI	HYPOTHETICAL 28.2 KD P	3.18e+00
16	64	29.2	357	1	FL3H_MATIN	NARINGENIN, 2-OXOGLUTAR	3.18e+00
17	64	29.2	369	1	FL3H_PETHY	NARINGENIN, 2-OXOGLUTAR	3.18e+00
18	64	29.2	577	1	CCBS_OENBE	PROBABLE CYTOCHROME C	3.18e+00
19	64	29.2	630	1	FIG1_MOUSE	FIG-1 PROTEIN PRECURSO	3.18e+00
20	64	29.2	846	1	PAC_ECOLI	PENICILLIN ACYLASE PRE	3.18e+00
21	63	28.8	295	1	UNG_MOUSE	URACIL-DNA GLYCOSYLASE	4.70e+00
22	63	28.8	396	1	BMP2_DAMDA	BONE MORPHOGENETIC PRO	4.70e+00
23	63	28.8	475	1	OPCA_SYNT3	PUTATIVE OXPPCYCLE PRO	4.70e+00

RESULT	ID	ECLC_BOVIN	STANDARD	PR	903 AA	ALIGNMENTS
AC	P54281					
DT	01-OCT-1996	(REL. 34, CREATED)				
DT	01-OCT-1996	(REL. 34, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)				
DE	EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE CHANNEL).					
OS	BOS TAURUS (BOVINE).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;					
OC	ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=TRACHEA;					
RX	MEDLINE; 96125078.					
RA	CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,					
RA	ARRATE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;					
RT	"Cloning of an epithelial chloride channel from bovine trachea.";					
RL	J. BIOL. CHEM. 270:31016-31026(1995).					
CC	-1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;					
CC	MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND TRANSEPITHELIAL TRANSPORT					
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).					
CC	-1- TISSUE SPECIFICITY: TRACHEA.					
CC	-1- PTM: PHOSPHORYLATED BY CAM-KINASE II.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	EMBL; U36445; G1184066; .					
DR	IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL; CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.					
KW	TRANSNEM 7 27 POTENTIAL.					
FT	TRANSNEM 331 351 POTENTIAL.					
FT	TRANSNEM 617 637 POTENTIAL.					
FT	TRANSNEM 883 903 POTENTIAL.					
FT	CARBOHYD 75 75 POTENTIAL.					
FT	CARBOHYD 278 278 POTENTIAL.					
FT	CARBOHYD 360 360 POTENTIAL.					
FT	CARBOHYD 372 372 POTENTIAL.					
FT	CARBOHYD 504 504 POTENTIAL.					

```
FT CARBOHYD 515 515 POTENTIAL.
FT CARBOHYD 688 688 POTENTIAL.
FT CARBOHYD 811 811 POTENTIAL.
FT CARBOHYD 816 816 POTENTIAL.
FT CARBOHYD 842 842 POTENTIAL.
FT CARBOHYD 857 857 POTENTIAL.
SQ SEQUENCE 903 AA; 100305 MW; 65E6EC1C CRC32;

Query Match 35.6%; Score 78; DB 1; Length 903;
Best Local Similarity 61.1%; Pred. No. 8.54e-03;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Db 522 TWTIKPEILLQDPKGGK 539
    ||| 1:||||| 1:|
QY 1 TWTTPPQILLWDPGQK 18

RESULT 2
ID AT12_HSV1F STANDARD; PRT: 715 AA.
AC P08314;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE ALPHA TRANS-INDUCING FACTOR 77 KD PROTEIN.
GN DL46.
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN F).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE: 87141362.
RA MCKNIGHT J.L.C., PELLET P.E., JENKINS F.J., ROIZMAN B.:
RT "Characterization and nucleotide sequence of two herpes simplex virus
RT 1 genes whose products modulate alpha-trans-inducing factor-dependent
RT activation of alpha genes.";
RL J. VIROL. 61:992-1003(1987).
CC -!- FUNCTION: MODULATE ALPHA TRANS-INDUCING FACTOR-DEPENDENT
CC ACTIVATION OF ALPHA GENES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15621; G330058; -.
CC FIR: B26133; TNBE77.
CC KW TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR.
CC SEQUENCE 715 AA; 77357 MW; F61478FC CRC32;

Query Match 33.3%; Score 73; DB 1; Length 715;
Best Local Similarity 37.0%; Pred. No. 7.70e-02;
Matches 10; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

Db 431 TWSAGPPDDMARGPGGHRAGGTGK 457
    ||| 1:|:|:|:| 1:|
QY 1 TWTTPPQILLWDPGQKGFVDK 26

RESULT 3
ID BMP2_CHICK STANDARD; PRT: 353 AA.
AC Q90751;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BONE MORPHOGENETIC PROTEIN 2 PRECURSOR (BMP-2) (FRAGMENT).
GN BMP2 OR BMP-2.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
[1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN;
RX MEDLINE: 94163974.
RA FRANCIS P.H., RICHARDSON M.K., BRICKELL P.M., TICKLE C.:
RT "Bone morphogenetic proteins and a signalling pathway that controls
RT patterning in the developing chick limb.";
RL DEVELOPMENT 120:209-218(1994).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X75914; G472928; -.
CC PROSITE: PS00250; TGF_BETA; 1.
CC PFAM: PF00019; TGF-beta; 1.
CC PFAM: PF00688; TGFb_propeptide; 1.
CC HSP: P18075; BMP.
CC GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.
KW NON_TER 1
FT PROPEP <1 239 BY SIMILARITY.
FT CHAIN 240 353 BONE MORPHOGENETIC PROTEIN 2.
FT DISULFID 253 318 BY SIMILARITY.
FT DISULFID 282 350 BY SIMILARITY.
FT DISULFID 286 352 BY SIMILARITY.
FT DISULFID 317 317 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 91 91 POTENTIAL.
FT CARBOHYD 121 121 POTENTIAL.
FT CARBOHYD 157 157 POTENTIAL.
FT CARBOHYD 295 295 POTENTIAL.
SQ SEQUENCE 353 AA; 40347 MW; 80C76C6E CRC32;

Query Match 32.4%; Score 71; DB 1; Length 353;
Best Local Similarity 47.4%; Pred. No. 1.81e-01;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Db 168 PAVLRWIAHQGNHGFVVE 186
    |:|:|:| 1:|
QY 7 PQILLWDPGQKQGGFVD 25

RESULT 4
ID BMP2_RAT STANDARD; PRT: 393 AA.
AC P49001;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE BONE MORPHOGENETIC PROTEIN 2 PRECURSOR (BMP-2) (BMP-2A).
GN BMP2 OR BMP-2.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-BONE;
RA FENG J.O., CHEN D., FENG M., HARRIS M.A., MUNDY G.R., HARRIS S.E.;
RL SUBMITTED (SEP-1993) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- TISSUE SPECIFICITY: FEMUR, CALVARIA, TRACHEA, LUNG AND OVARY.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z25868; G397951; -  
CC PIR: A34201; A34201.  
CC MGI: 88177; BMP2.  
CC PROSITE: PS00250; TGF-BETA; 1.  
CC PFAM: PF00019; TGF-beta; 1.  
CC PFAM: PF00688; TGF-beta; 1.  
CC HSSP: P18075; 1BMP.  
KW SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.  
FT SIGNAL 1 ?  
FT PROPEP 1 ?  
FT CHAIN 280 ?  
FT CHAIN 280 ?  
FT DISULFID 293 358 BONE MORPHOGENETIC PROTEIN 2.  
FT DISULFID 322 390 BY SIMILARITY.  
FT DISULFID 326 392 BY SIMILARITY.  
FT DISULFID 357 357 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 133 133 POTENTIAL.  
FT CARBOHYD 161 161 POTENTIAL.  
FT CARBOHYD 197 197 POTENTIAL.  
FT CARBOHYD 335 335 POTENTIAL.  
SQ SEQUENCE 393 AA; 44383 MW; 582689AF CRC32;  
  
Query Match 32.08; Score 70; DB 1; Length 393;  
Best Local Similarity 36.88; Pred. No. 2.76e-01;  
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
  
Db 208 PAVMRWTAQHTNHGFVVE 226  
QY 7 PQILLWDPSGQKGGFVVD 25  
  
RESULT 5  
ID BMP2\_MOUSE STANDARD; PRT; 394 AA.  
AC P21274;  
DT 01-MAY-1991 (REL. 18, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE BONE MORPHOGENETIC PROTEIN 2 PRECURSOR (BMP-2) (BMP-2A).  
GN BMP2 OR BMP-2  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 94289485.  
RA FENG J.Q., HARRIS M.A., GHOSH-CHOUDHURY N., FENG M., MUNDY G.R.,  
RA HARRIS S.E.;  
RT "Structure and sequence of mouse bone morphogenetic protein-2 gene  
(BMP-2): comparison of the structures and promoter regions of BMP-2  
and BMP-4 genes";  
RL BIOCHIM. BIOPHYS. ACTA 1218:221-224(1994).  
RN [2]  
RP SEQUENCE OF 1-351 FROM N.A.  
RX MEDLINE: 90228966.  
RA DICKINSON M.E., KOBIRIN M.S., SILAN C.M., KINGSLEY D.M., JUSTICE M.J.,  
RA MILLER D.A., CECI J.D., LOCK L.F., LEE A., BUCHBERG A.M.,  
RA SIRACUSA L.D., LYONS K.M., DERYNCK R., HOGAN B.L.M., COPELAND N.G.,  
RA JENKINS N.A.;  
RT "Chromosomal localization of seven members of the murine TGF-beta  
superfamily suggests close linkage to several morphogenetic mutant  
loci";  
RL GENOMICS 6:505-520(1990).  
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.  
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.  
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L25602; G409362; -  
CC PIR: A34201; A34201.  
CC MGI: 88177; BMP2.  
CC PROSITE: PS00250; TGF-BETA; 1.  
CC PFAM: PF00019; TGF-beta; 1.  
CC PFAM: PF00688; TGF-beta; 1.  
CC HSSP: P18075; 1BMP.  
KW SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.  
FT SIGNAL 1 ?  
FT PROPEP 1 ?  
FT CHAIN 280 ?  
FT CHAIN 280 ?  
FT DISULFID 293 358 BONE MORPHOGENETIC PROTEIN 2.  
FT DISULFID 322 390 BY SIMILARITY.  
FT DISULFID 326 392 BY SIMILARITY.  
FT DISULFID 357 357 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 133 133 POTENTIAL.  
FT CARBOHYD 161 161 POTENTIAL.  
FT CARBOHYD 197 197 POTENTIAL.  
FT CARBOHYD 335 335 POTENTIAL.  
SQ SEQUENCE 394 AA; 44514 MW; B055A5A9 CRC32;  
  
Query Match 31.5%; Score 69; DB 1; Length 394;  
Best Local Similarity 36.88; Pred. No. 4.19e-01;  
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
  
Db 209 PAVMRWTTQHTNHGFVVE 227  
QY 7 PQILLWDPSGQKGGFVVD 25  
  
RESULT 6  
ID UNG\_PSEDO STANDARD; PRT; 217 AA.  
AC P29950;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE URACIL-DNA GLYCOSYLASE (EC 3.2.2.-) (UDG) (FRAGMENT).  
GN UNG.  
OS PSEUDOMONAS DENITRIFICANS.  
OC BACTERIA; PROTEOBACTERIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-SC510;  
RX MEDLINE: 91008976.  
RA CROUZET J., CAMERON B., CAUCHOIS L., RIGAUULT S., ROUYEZ M.-C.,  
RA BLANCHE F., THIBAUT D., DEBUSSCHE L.;  
RT "Genetic and sequence analysis of an 8.7-kilobase Pseudomonas  
denitrificans fragment carrying eight genes involved in  
transformation of precorrin-2 to cobyrinic acid";  
RL J. BACTERIOL. 172:5980-5990(1990).  
RN [2]  
RP IDENTIFICATION, AND CORRECTION OF A PROBABLE FRAMESHIFT.  
RA AASLAND R.;  
RL UNPUBLISHED OBSERVATIONS (FEB-1993).  
CC -!- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE  
CC AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA  
CC POLYMERASE OR DUE TO DEAMINATION OF CYTOSINE (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE URACIL-DNA GLYCOSYLASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M59301; -; NOT\_ANNOTATED\_CDS.  
CC PIR: I36145; I36145.

```
DR PROSITE; PS00130; U_DNA_GLYCOSYLASE; 1.
DR PFAM; PF00315; UNG; 1.
DR HSSP; P13051; IAKZ.
KW DNA REPAIR; HYDROLASE; GLYCOSIDASE.
FT NON_TER 1
FT ACT_SITE 51 GENERAL BASE (BY SIMILARITY).
SQ SEQUENCE 217 AA; 24314 MW; 05D9E603 CRC32;

Query Match 31.1%; Score 68; DB 1; Length 217;
Best Local Similarity 36.8%; Pred. No. 6.34e-01;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 143 HPVFMWGSYAQKAAVF 161
QY 5 QPQILLWDPGKQGQGV 23

RESULT 7
ID YBIR_ECOLI STANDARD; PRT; 372 AA.
AC P75788;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 41.2 KD PROTEIN IM OMPX-MOEB INTERGENIC REGION.
GN YBIR.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE; 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL SCIENCE 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: TO M.JANNASCHII MJ0456.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000184; G187039; -.
DR ECOGENE; EG13323; YBIR.
KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 248 268 POTENTIAL.
FT TRANSMEM 284 304 POTENTIAL.
FT TRANSMEM 310 330 POTENTIAL.
FT TRANSMEM 349 369 POTENTIAL.
SQ SEQUENCE 372 AA; 41168 MW; 3F83745E CRC32;

Query Match 31.1%; Score 68; DB 1; Length 372;
Best Local Similarity 38.1%; Pred. No. 6.34e-01;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Db 143 PONILWGRSGLSFAGFAQM 163
QY 6 PPOILLWDPGKQGQGVVDK 26

DR PROSITE; PS00130; U_DNA_GLYCOSYLASE; 1.
DR PFAM; PF00315; UNG; 1.
DR HSSP; P13051; IAKZ.
KW DNA REPAIR; HYDROLASE; GLYCOSIDASE.
FT NON_TER 1
FT ACT_SITE 51 GENERAL BASE (BY SIMILARITY).
SQ SEQUENCE 217 AA; 24314 MW; 05D9E603 CRC32;

Query Match 31.1%; Score 68; DB 1; Length 217;
Best Local Similarity 36.8%; Pred. No. 6.34e-01;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 143 HPVFMWGSYAQKAAVF 161
QY 5 QPQILLWDPGKQGQGV 23

RESULT 8
ID BMP2_HUMAN STANDARD; PRT; 396 AA.
AC P12643;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE BONE MORPHOGENETIC PROTEIN 2 PRECURSOR (BMP-2) (BMP-2A).
GN BMP2 OR BMP2A.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89072730.
RA WOZNEY J.M., ROSEN V., CELESTE A.J., MITSOCK L.M., WHITTERS M.J.,
RA KRIZ R.W., HEWICK R.M., WANG E.A.;
RT "Novel regulators of bone formation: molecular clones and
RT activities".
RL SCIENCE 242:1528-1534(1988).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- TISSUE SPECIFICITY: PARTICULARLY ABUNDANT IN LUNG, SPLEEN AND
CC COLON AND IN LOW BUT SIGNIFICANT LEVELS IN HEART, BRAIN, PLACENTA,
CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS, PROSTATE, OVARY AND
CC SMALL INTESTINE.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22489; G179502; -.
DR PIR; B37278; B37278.
DR MIN; 112261; -.
DR PROSITE; PS00250; TGF-BETA; 1.
DR PFAM; PF00019; TGF-beta; 1.
DR PFAM; PF00688; TGFb_propeptide; 1.
DR HSSP; P18075; IBMP.
KW SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.
FT SIGNAL 1 ? 282 POTENTIAL.
FT PROPEP ? 282
FT CHAIN 283 396 BONE MORPHOGENETIC PROTEIN 2.
FT DISULFID 296 361 BY SIMILARITY.
FT DISULFID 325 393 BY SIMILARITY.
FT DISULFID 329 395 BY SIMILARITY.
FT DISULFID 360 360 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 135 135 POTENTIAL.
FT CARBOHYD 163 163 POTENTIAL.
FT CARBOHYD 164 164 POTENTIAL.
FT CARBOHYD 200 200 POTENTIAL.
SQ SEQUENCE 396 AA; 44702 MW; 182782C0 CRC32;

Query Match 30.6%; Score 67; DB 1; Length 396;
Best Local Similarity 36.8%; Pred. No. 9.55e-01;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 211 PAVMRTAQGHANHFVVE 229
QY 7 POILLWDPGKQGQGVVD 25

RESULT 9
ID GSP1_XANCP STANDARD; PRT; 138 AA.
AC P31738;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE GENERAL SECRETION PATHWAY PROTEIN I PRECURSOR.
GN XPSI OR PEPI.
```





RESULT	13	FAK1_HUMAN	STANDARD;	PRT; 1052 AA.
ID		FAK1_Q14291;		
AC		Q05397; Q14291;		
DT	01-FEB-1994	(REL. 28, CREATED)		
DT	15-JUL-1998	(REL. 36, LAST SEQUENCE UPDATE)		
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)		
DE	FOCAL ADHESION KINASE 1 (EC 2.7.1.112) (FADK 1) (PPI25FAK) (PROTEIN-			



Matches	9;	Conservative	7;	Mismatches	6;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

```

Db 575 APQIILREADNYKQVGAEVEKE 596
QV 6 PPOILLWDPSGOKOGGFVVVDKN 27

```

RESULT 15

ID	YNFP_ECOLI	STANDARD;	PRT;	263 AA.
AC	P75981;			
DT	15-DEC-1998	(REL. 37, CREATED)		
DT	15-DEC-1998	(REL. 37, LAST SEQUENCE UPDATE)		
DT	15-DEC-1998	(REL. 37, LAST ANNOTATION UPDATE)		
DE	HYPOTHETICAL 28.2 KD PROTEIN IN INTE-PIN INTERGENIC REGION.			
GN	YNFP.			
OS	ESCHERICHIA COLI.			
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;			
QC	ESCHERICHIA.			

[1] SEQUENCE FROM N.A.  
RP STRAIN=K12 / MG1655;  
RX MEDLINE: 97426617.  
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
RX RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
RX MAU B., SHAO Y.:  
RT "The complete genome sequence of *Escherichia coli* K-12";  
RL SCIENCE 277:1453-1474(1997).  
CC -1- SIMILARITY: TO B.SUBTILIS XKDT AND YQBT AND TO H.INFLENZAE HI1520.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR ENBL; AE000214; G1787398; -.  
DR ECGENE; EG14254; YMFP.  
DR ECOTHEP; EG14254; YMFP.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 263 AA; 28232 MW; BC3D8190 CRC32;

SEQUENCE 263 AA; 28232 MW; BC3D8190 CRC32;

Query Match 29.28: Score 64: DB 1: Length 263:

Best Local Similarity 33.38; Pred. No. 3.18e+00;

Matches 8; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Db 92 TWRARVIERYYWTPGGADGDYV 115

—  
..  
—  
..  
—  
—  
..  
..  
—  
—

Qy 1 TWTTQPPIILLWDPGQKQGGFVV 24

Search completed: Mon Aug 23 13:10:15 1999

Job time : 11 secs.

2  
2  
3  
3  
4  
4  
-  
0  
3  
4  
0  
2  
0  
0

\*\*\*\*\*  
W P E R E L L (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:10:33 1999; MasPar time 12.31 Seconds  
Tabular output not generated.  
128.612 Million cell updates/sec

Title: >US-09-049-696-45  
Description: (1-29) from US09049696.pep  
Perfect Score: 219  
Sequence: 1 TWTTPPQILLWDPGQKGQGFVVDKNTK 29

Scoring table: PAM 150  
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sprembl9  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp\_mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 30.154; Variance 46.515; scale 0.648

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	148	67.6	913	11	O88826	8.62e-17
2	78	35.6	794	6	O18742	3.83e-02
3	78	35.6	820	6	O18743	3.83e-02
4	78	35.6	905	6	O18741	3.83e-02
5	77	35.2	366	10	O23825	5.80e-02
6	75	34.2	144	1	O30283	1.32e-01
7	69	31.5	625	10	O96466	1.45e+00
8	67	30.6	395	6	O46564	3.12e+00
9	67	30.6	862	2	O86144	3.12e+00
10	67	30.6	1229	5	O20137	3.12e+00
11	66	30.1	357	10	O42927	4.56e+00
12	66	30.1	357	10	O40344	4.56e+00
13	66	30.1	475	13	O93531	4.56e+00
14	66	30.1	796	2	O69908	4.56e+00
15	66	30.1	901	11	O88860	4.56e+00
16	66	30.1	1068	13	O91563	4.56e+00
17	65	29.7	315	10	O81535	6.64e+00
18	65	29.7	369	10	O40754	6.64e+00
19	65	29.7	507	10	O22233	6.64e+00
20	65	29.7	614	13	O90484	6.64e+00

21	65	29.7	642	14	O90184	HOMOLOGIES WITH ORF 46	6.64e+00
22	65	29.7	1164	4	O92574	HAMARIN (MYELOBLAST K	6.64e+00
23	64	29.2	233	1	O52013	ORF H1320.	9.63e+00
24	64	29.2	298	5	O62415	Y22FSA.4 PROTEIN.	9.63e+00
25	64	29.2	366	10	O22530	FLAVANONE 3BETA-HYDROX	9.63e+00
26	64	29.2	367	10	O96472	FLAVANONE 3-HYDROXYLAS	9.63e+00
27	64	29.2	374	10	O43741	NARINGENIN 3-DIOXYGENA	9.63e+00
28	64	29.2	381	8	P92585	CC11.	9.63e+00
29	64	29.2	382	8	O31706	CC11-LIKE PROTEIN.	9.63e+00
30	64	29.2	636	2	P74536	HYPOTHETICAL 71.6 KD P	9.63e+00
31	64	29.2	710	3	O00210	CHROMOSOME SEGREGATION	9.63e+00
32	64	29.2	1022	5	O17943	K09F5.3 PROTEIN.	9.63e+00
33	64	29.2	1026	5	O18987	D2045.1 PROTEIN.	9.63e+00
34	64	29.2	1825	5	O61210	H19M22.1 PROTEIN (FRAG	9.63e+00
35	63	28.8	306	11	P97285	URACIL-DNA GLYCOSYLASE	1.39e+01
36	63	28.8	434	2	P75037	SIMILAR TO GENBANK ACC	1.39e+01
37	63	28.8	686	11	P97860	BIG TAU.	1.39e+01
38	63	28.8	707	11	P97860	LEUCINE-RICH REPEAT PR	1.39e+01
39	63	28.8	915	2	O66044	VAG8.	1.39e+01
40	63	28.8	1005	2	O45703	TRANSPOSASE.	1.39e+01
41	63	28.8	1199	14	O83859	136.6KD PROTEIN.	1.39e+01
42	63	28.8	1512	2	O85324	HYPOTHETICAL 161.8 KD	1.39e+01
43	63	28.8	5825	10	O82731	ORFA+B.	1.39e+01
44	62	28.3	358	10	O38877	FLAVANONE 3-HYDROXYLAS	2.00e+01
45	62	28.3	1526	5	O18023	C25G4.10 PROTEIN.	2.00e+01

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	913 AA.
ID	O88826			
AC	O88826			
DT	01-NOV-1998 (TREMREL. 08, CREATED)			
DT	01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)			
DE	G0B-5 PROTEIN.			
GN	G0B-5.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			
OC	SCIUROGNATHI; MURIDAE; MURINAE; MUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=INTESTINE;			
RA	KOMIYA T., TANIGAWA Y., HIROHASHI S.;			
RT	"Cloning of a gene, gob-5, which is expressed in intestinal goblet cells in mice."			
RL	SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: AB017156; D1034712; "			
SQ	SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;			

Query Match 67.6%; Score 148; DB 11; Length 913;  
Best Local Similarity 62.1%; Pred. No. 8.62e-17;  
Matches 18; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db	522	TWTHPTPIWDPSPGVEQNGFILDITTK	550
QY	1	TWTTPPQILLWDPGQKGQGFVVDKNTK	29

RESULT	2	PRELIMINARY;	PRT;	794 AA.
ID	O18742			
AC	O18742			
DT	01-JAN-1998 (TREMREL. 05, CREATED)			
DT	01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)			
DT	01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)			
DE	LU-ECAM-1.			
OS	BOS TAURUS (BOVINE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RA	ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,			

RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF001262; G2623765; -  
SQ SEQUENCE 794 AA; 88509 MW; 2D450EEB CRC32;

Query Match 35.6%; Score 78; DB 6; Length 794;  
Best Local Similarity 61.1%; Pred. No. 3.83e-02;  
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIQPEIVLQDPKGGK 540  
||| | | | | | | | | |  
QY 1 TWTTPPQILLWDPGQK 18

RESULT 3

ID O18743 PRELIMINARY; PRT; 820 AA.  
AC O18743;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
DE LU-ECAM-1.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF001263; G2623767; -  
SQ SEQUENCE 820 AA; 91464 MW; 171ED21C CRC32;

Query Match 35.6%; Score 78; DB 6; Length 820;  
Best Local Similarity 61.1%; Pred. No. 3.83e-02;  
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIQPEIVLQDPKGGK 540  
||| | | | | | | | | |  
QY 1 TWTTPPQILLWDPGQK 18

RESULT 4

ID O18741 PRELIMINARY; PRT; 905 AA.  
AC O18741;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE LU-ECAM-1.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF001261; G2623763; -  
DR PFAM; PF00041; fn3; 1.  
SQ SEQUENCE 905 AA; 101005 MW; 86E65459 CRC32;

Query Match 35.6%; Score 78; DB 6; Length 905;  
Best Local Similarity 61.1%; Pred. No. 3.83e-02;  
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 523 TWTIQPEIVLQDPKGGK 540  
||| | | | | | | | | |  
QY 1 TWTTPPQILLWDPGQK 18

RESULT 5

ID O23825 PRELIMINARY; PRT; 366 AA.

AC O23825;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE 2-OXOGULUTARATE 3-DIOXYGENASE (EC 1.14.11.9).  
GN F3H-A.  
OS IPOMOEA NIL (JAPANESE MORNING GLORY).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC ASTERIDAE; SOLANACEAE; CONVULVACEAE; IPOMOEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KK/ZSK-2; TISSUE=YOUNG BUDS;  
RX MEDLINE; 97468507.  
RA HOSHINO A., ABE Y., SAITO N., INAGAKI Y., IIDA S.;  
RT "The gene encoding flavanone 3-hydroxylase is expressed normally in  
RT the pale yellow flowers of the Japanese morning glory carrying the  
RT speckled mutation which produce neither flavonol nor anthocyanin but  
RT accumulate chalcone, aurone and flavanone.";  
RL PLANT CELL PHYSIOL. 38:970-974 (1997).  
DR EMBL; D83041; D1022755; -  
DR PFAM; PF00671; Fe\_Asc-Oxidore; 1.  
KW DIOXYGENASE; OXIDOREDUCTASE.  
SQ SEQUENCE 366 AA; 41001 MW; C0F12E52 CRC32;

Query Match 35.2%; Score 77; DB 10; Length 366;  
Best Local Similarity 50.0%; Pred. No. 5.80e-02;  
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Db 98 PEEKLLFDMSGGKGGFTVSSH 119  
| : | | | | | | | | | | :  
QY 6 PPQILLWDPGQKGGFVVDKN 27

RESULT 6

ID O30283 PRELIMINARY; PRT; 144 AA.  
AC O30283;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
DE CONSERVED HYPOTHETICAL PROTEIN.  
GN AF2387.  
OS ARCHAEoglobus fulgidus.  
OC ARCHAEA; EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobaceae;  
OC ARCHAEoglobus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE; 98049343.  
RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,  
RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,  
RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,  
RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
RA OVERBEK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,  
RA CORTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
RA VENTER J.C.;  
RT "The complete genome sequence of the hyperthermophilic,  
RT sulphate-reducing archaeon Archaeoglobus fulgidus.";  
RL NATURE 390:364-370 (1997).  
DR EMBL; AE001111; G2650702; -  
DR TIGR; AF2387; -  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 144 AA; 16212 MW; BDD8BADE CRC32;

Query Match 34.2%; Score 75; DB 1; Length 144;  
Best Local Similarity 46.7%; Pred. No. 1.32e-01;  
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 49 LLLWASRRRRGGFV 63

```
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 210 PAVMRWTAQGHANHGFFVE 228
QY 7 POILLWDPGKQGGFVVD 25

RESULT 9
ID 086144 PRELIMINARY; PRT; 862 AA.
AC 086144;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE SUCROSE FRUCTAN 6-FRUCTOSYL TRANSFERASE.
OS HORDEUM VULGARE (BARLEY).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
OC POACEAE; HORDEUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EXPRESS; TISSUE=LEAF;
RX SPRENGER N.; BORTLIK K.; BRANDT A.; BOLLER T.; WIEMKEN A.;
RT "Purification, cloning, and functional expression of sucrose:fructan
6-fructosyltransferase, a key enzyme of fructan synthesis in
barley.";
RL PROC. NATL. ACAD. SCI. U.S.A. 92:11652-11656(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EXPRESS; TISSUE=LEAF;
RX SPRENGER N.;
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EXPRESS; TISSUE=LEAF;
RX SPRENGER N.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X83233; E266642; -.
DR PFAM; PF00251; glycosyl_hydro2; 1.
DR MENDEL; 8477; HORVU; 1002; 1.
KW TRANSFERASE
SQ SEQUENCE 625 AA; 69410 MW; 9B8E15CF CRC32;

Query Match 31.5%; Score 69; DB 10; Length 625;
Best Local Similarity 40.0%; Pred. No. 1.45e+00;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 200 WTKHPANVIVSPG 214
QY 2 WTTQPPQILLWDPG 16

RESULT 8
ID 046564 PRELIMINARY; PRT; 395 AA.
AC 046564;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE BONE MORPHOGENETIC PROTEIN 2.
GN BMP-2.
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE;
RX WAN X.L.; SEARS J.; CHEN S.; SEARS M.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
DR EMBL; AF041421; G273388; -.
DR PROSITE; PS00250; TGF-BETA; 1.
KW GLYCOPROTEIN.
SQ SEQUENCE 395 AA; 44664 MW; 81F38F1E CRC32;

Query Match 30.6%; Score 67; DB 6; Length 395;
Best Local Similarity 36.8%; Pred. No. 3.12e+00;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 210 PAVMRWTAQGHANHGFFVE 228
QY 7 POILLWDPGKQGGFVVD 25

RESULT 9
ID 086144 PRELIMINARY; PRT; 862 AA.
AC 086144;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE FLAGELLAR HOOK SUBUNIT PROTEIN.
GN FLGE.
OS CAMPYLOBACTER JEJUNI.
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
OC CAMPYLOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LIO 36;
RX MEDLINE; 98324984.
RA LUENEBERG E.; GLENN-CALVO E.; HARTMANN M.; BAER W.; FROSC M.;
RT "The central, surface-exposed region of the flagellar hook protein
J. BACTERIOL. 180:3711-3714(1998).
RL EMBL; AJ002074; E1308747; -.
KW FLAGELLA.
SQ SEQUENCE 862 AA; 91550 MW; 8101430F CRC32;

Query Match 30.6%; Score 67; DB 2; Length 862;
Best Local Similarity 50.0%; Pred. No. 3.12e+00;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 380 WDPALMGQGGFTFNQNT 397
QY 11 LWDPGKQGGFVVDKNT 28

RESULT 10
ID Q20137 PRELIMINARY; PRT; 1229 AA.
AC Q20137;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE C09D8.1 (FRAGMENT).
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA SWINBURNE J.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R.; AINSCOUGH R.; ANDERSON K.; BAYNES C.; BERKS M.;
RA BONFIELD J.; BURTON J.; CONNELL M.; COPSEY T.; COOPER J.; COULSON A.;
RA CRAXTON M.; DEAR S.; DU Z.; DURBIN R.; FAVELLO A.; FULTON L.;
RA GARDNER A.; GREEN P.; HAWKINS T.; HILLIER L.; JIER M.; JOHNSTON L.;
RA JONES M.; KERSHAW J.; KIRSTEN J.; LAISTER N.; LATREILLE P.;
RA LIGHTNING J.; LLOYD C.; MCMURRAY A.; MORTIMORE B.; O'CALLAGHAN M.;
RA PARSONS J.; PERCY C.; RIFKEN L.; ROOPRA A.; SAUNDERS D.; SHOWNKEEN R.;
RA SWALDON N.; SMITH A.; SONNHAMMER E.; STADEN R.; SULTON J.;
RA THIERRY-MIEG J.; THOMAS K.; VAUDIN M.; VAUGHAN K.; WATERSTON R.;
RA WATSON A.; WEINSTOCK L.; WILKINSON-SPROAT J.; WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL NATURE 368:32-38(1994).
DR EMBL; Z49938; G895771; -.
DR PFAM; PF00041; fn3; 1.
DR PFAM; PF00102; Y_phosphatase; 2.
RN NON_TER 1 1
```

SQ SEQUENCE 1229 AA; 138963 MW; 0FC35C5A CRC32;  
 Query Match 30.6%; Score 67; DB 5; Length 1229;  
 Best Local Similarity 28.6%; Pred. No. 3.12e+00;  
 Matches 8; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Db 218 WKTNSVOLTWEYNGPRNGFYVNHGTR 245  
 I : I : I : I : I : I : I :  
 Qy 2 WTTQPPQILLMDPSGQKGQGFVVDKNTK 29

RESULT 11  
 ID Q42927 PRELIMINARY; PRT; 357 AA.  
 AC Q42927;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE NARINGENIN 3-DIOXYGENASE (EC 1.14.11.9)  
 DE (NARINGENIN; 2-OXOGLUTARATE 3-DIOXYGENASE) (FLAVANONE 3-HYDROXYLASE).  
 GN F3H1.  
 OS MEDICAGO SATIVA (ALFALFA).  
 OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC FABALES; FABACEAE; PAPILIONOIDEAE; MEDICAGO.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CV. NAGYSZENASI;  
 RX MEDLINE: 96128019.  
 RA CHARRIER B., CORONADO C., KONDOROSI A., RATET P.;  
 RT "Molecular characterization and expression of alfalfa (Medicago sativa L.) flavanone-3-hydroxylase and dihydroflavonol-4-reductase encoding genes.";  
 RT PLANT MOL. BIOL. 29:773-786(1995).  
 CC -!- CATALYTIC ACTIVITY: NARINGENIN + 2-OXOGLUTARATE + O(2) =  
 CC DHYDROKAEPPFEROL + SUCCINATE + CO(2).  
 CC -!- COFACTOR: IRON; ASCORBATE.  
 CR EMBL: X78994; G475959; -;  
 DR PFAM: PF00671; Fe\_Asc\_Oxidored; 1.  
 DR MENDEL: 9062; MEDsa:1186.2.  
 KW DIOXYGENASE; OXIDOREDUCTASE.  
 SQ SEQUENCE 357 AA; 40607 MW; 5AAB8D92 CRC32;

Query Match 30.1%; Score 66; DB 10; Length 357;  
 Best Local Similarity 50.0%; Pred. No. 4.56e+00;  
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Db 98 PEEKLRFDMSGGKKGFFVSSH 119  
 I : I : I : I : I : I : I :  
 Qy 6 PPOILLMDPSGQKGQGFVVDKNTK 27

RESULT 12  
 ID Q40344 PRELIMINARY; PRT; 357 AA.  
 AC Q40344;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE FLAVANONE-3-HYDROXYLASE.  
 GN F3H2.  
 OS MEDICAGO SATIVA (ALFALFA).  
 OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
 OC FABALES; FABACEAE; PAPILIONOIDEAE; MEDICAGO.  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CV. NAGYSZENASI;  
 RX MEDLINE: 96311005.  
 RA CHARRIER B., LEROUX C., KONDOROSI A., RATET P.;  
 RT "The expression pattern of alfalfa flavanone 3-hydroxylase promoter-gus fusion in Nicotiana benthamiana correlates with the presence of flavonoids detected in situ.";  
 RT PLANT MOL. BIOL. 30:1153-1168(1996).  
 DR EMBL: X81812; G550391. -;  
 SQ SEQUENCE 357 AA; 40607 MW; 5AAB8D92 CRC32;





...is Page Blank (uspto)



90	5	17.2	687 34	W71183	Mutant human lactofer	2.62e+02	163	5	17.2	1403 3	R10890	lktA::lacZ fusion pro	2.62e+02
91	5	17.2	688 34	W71182	Mutant human lactofer	2.62e+02	164	5	17.2	3724 26	W23718	Platenolide synthase	2.62e+02
92	5	17.2	689 34	W71181	Mutant human lactofer	2.62e+02	165	5	17.2	3724 26	W22608	Platenolide synthase	2.62e+02
93	5	17.2	690 34	W71180	Mutant human lactofer	2.62e+02	166	4	13.8	7 24	W25919	Beta-2-microglobulin	2.47e+03
94	5	17.2	693 14	R85146	Lactoferrin protein.	2.62e+02	167	4	13.8	8 7	R35831	Hepatitis C virus (HC	2.47e+03
95	5	17.2	705 4	R22423	Human lactoferrin ded	2.62e+02	168	4	13.8	9 11	R61061	Dynorphin-like polype	2.47e+03
96	5	17.2	708 4	R22424	Human lactoferrin ded	2.62e+02	169	4	13.8	10 12	R63347	Peptide fragment of v	2.47e+03
97	5	17.2	709 12	R45199	Human lactoferrin.	2.62e+02	170	4	13.8	10 4	P40448	Sequence of peptide w	2.47e+03
98	5	17.2	709 9	R45198	Human lactoferrin.	2.62e+02	171	4	13.8	10 26	W26228	Fusion protein spacer	2.47e+03
99	5	17.2	709 30	W53879	Human lactoferrin.	2.62e+02	172	4	13.8	11 36	W57243	aPL immunoreactive pe	2.47e+03
100	5	17.2	709 23	W21695	Human lactoferrin enc	2.62e+02	173	4	13.8	12 36	W77393	Gibbon ape leukemia v	2.47e+03
101	5	17.2	709 3	R12583	Lactoferrin.	2.62e+02	174	4	13.8	13 13	R66881	Agonist peptide capab	2.47e+03
102	5	17.2	711 38	W85021	Human lactoferrin (hL	2.62e+02	175	4	13.8	15 39	W73726	M. tuberculosis antig	2.47e+03
103	5	17.2	711 20	W09342	Human lactoferrin.	2.62e+02	176	4	13.8	15 17	R96960	Multi-drug resistance	2.47e+03
104	5	17.2	711 30	W53880	Bovine alphaS1-casein	2.62e+02	177	4	13.8	15 39	W38336	M. tuberculosis antig	2.47e+03
105	5	17.2	711 8	R43653	Lactoferrin.	2.62e+02	178	4	13.8	15 26	W38976	Peptide resembling an	2.47e+03
106	5	17.2	711 2	R08037	Human lactoferrin.	2.62e+02	179	4	13.8	16 17	R90961	Factor viii antigenic	2.47e+03
107	5	17.2	711 31	W57317	Human lactoferrin.	2.62e+02	180	4	13.8	16 39	W89185	Osif2/Cbfa1 peptide ep	2.47e+03
108	5	17.2	739 7	R38266	Sequence of a bovine	2.62e+02	181	4	13.8	17 38	W88861	Polypeptide fragment	2.47e+03
109	5	17.2	821 14	R05074	Arabidopsis CTR1 prot	2.62e+02	182	4	13.8	18 11	R59859	Type I collagen C-tel	2.47e+03
110	5	17.2	821 22	W17938	Constitutive triple r	2.62e+02	183	4	13.8	19 39	W85478	Mouse GP88 autocrine	2.47e+03
111	5	17.2	821 9	R46723	Arabidopsis thaliana	2.62e+02	184	4	13.8	20 7	R37301	Rabbit Muscle aldolas	2.47e+03
112	5	17.2	845 3	P50035	N-terminal sequence o	2.62e+02	185	4	13.8	20 18	R90974	HCV E1 peptide EI-49	2.47e+03
113	5	17.2	846 37	W79532	Human CLOCK protein.	2.62e+02	186	4	13.8	20 11	R59860	Type I collagen C-tel	2.47e+03
114	5	17.2	850 33	W62040	Protein isolated from	2.62e+02	187	4	13.8	20 30	W53625	Mouse bax protein res	2.47e+03
115	5	17.2	854 2	P70347	Varicella-zoster viru	2.62e+02	188	4	13.8	22 29	W33982	Dopamine transporter	2.47e+03
116	5	17.2	868 4	R22615	Varicella-zoster viru	2.62e+02	189	4	13.8	22 35	W69763	Proteolipid protein f	2.47e+03
117	5	17.2	876 35	W69349	Heparitinase protein.	2.62e+02	190	4	13.8	22 32	W84793	PLP 170-191.	2.47e+03
118	5	17.2	886 39	W89787	Staphylococcus aureus	2.62e+02	191	4	13.8	25 6	R31977	BCH-456.	2.47e+03
119	5	17.2	900 15	R88616	TNF-R p55IC-binding p	2.62e+02	192	4	13.8	25 12	R68761	Cytotoxic T lymphocyt	2.47e+03
120	5	17.2	907 14	R80144	EBV gp350/220.	2.62e+02	193	4	13.8	25 35	W66091	Hepatitis C virus p21	2.47e+03
121	5	17.2	908 21	P50073	Epstein-Barr virus (E	2.62e+02	194	4	13.8	26 17	R52285	Mouse heavy chain sur	2.47e+03
122	5	17.2	907 2	W15473	Human p100 protein.	2.62e+02	195	4	13.8	28 6	R31559	HTLV-II epitope, HTLV	2.47e+03
123	5	17.2	911 18	R77892	Bacterial transferrin	2.62e+02	196	4	13.8	30 39	W57827	Human secreted protei	2.47e+03
124	5	17.2	911 30	W54124	H. influenzae strain	2.62e+02	197	4	13.8	30 17	R85820	Peptide r111 from tne	2.47e+03
125	5	17.2	911 28	W08966	Amino acid sequence o	2.62e+02	198	4	13.8	32 6	R31557	HTLV-II epitope, HTLV	2.47e+03
126	5	17.2	912 28	W08961	Amino acid sequence o	2.62e+02	199	4	13.8	33 36	W74953	Human secreted protei	2.47e+03
127	5	17.2	912 30	W53044	H. influenzae strain	2.62e+02	200	4	13.8	34 6	R31563	HTLV-II epitope, HTLV	2.47e+03
128	5	17.2	912 17	R77886	Bacterial transferrin	2.62e+02	201	4	13.8	37 38	W88771	Polyptide fragment	2.47e+03
129	5	17.2	912 28	W08959	Amino acid sequence o	2.62e+02	202	4	13.8	38 38	W88547	Secreted protein enco	2.47e+03
130	5	17.2	912 18	R77888	Bacterial transferrin	2.62e+02	203	4	13.8	40 38	W88718	Secreted protein enco	2.47e+03
131	5	17.2	912 17	R96969	Transferrin receptor	2.62e+02	204	4	13.8	41 4	R21452	Peptide BCH-456 deriv	2.47e+03
132	5	17.2	912 30	W53046	H. influenzae strain	2.62e+02	205	4	13.8	53 10	R55159	CMP-NeuAc hydroxylas	2.47e+03
133	5	17.2	913 30	W53042	H. influenzae strain	2.62e+02	206	4	13.8	55 8	R44503	Sequence of N-termina	2.47e+03
134	5	17.2	913 28	W08957	Amino acid sequence o	2.62e+02	207	4	13.8	55 38	W88621	Secreted protein enco	2.47e+03
135	5	17.2	913 34	W68404	Aujeszky's disease vi	2.62e+02	208	4	13.8	56 38	W88781	Polyptide fragment	2.47e+03
136	5	17.2	913 18	R77884	Bacterial transferrin	2.62e+02	209	4	13.8	58 38	W88954	Sequence ID #654 from	2.47e+03
137	5	17.2	914 28	W08963	Amino acid sequence o	2.62e+02	210	4	13.8	58 39	W87540	Peptide derived from	2.47e+03
138	5	17.2	914 30	W53048	H. influenzae strain	2.62e+02	211	4	13.8	65 39	W88227	Human CD27 binding pr	2.47e+03
139	5	17.2	928 14	R77399	BHVI gi glycoprotein.	2.62e+02	212	4	13.8	72 31	W28171	Amino acid sequence o	2.47e+03
140	5	17.2	930 32	W5961	Human transient recep	2.62e+02	213	4	13.8	72 10	W58812	E. coli multiple anti	2.47e+03
141	5	17.2	930 23	W18061	Pasteurella haemolyti	2.62e+02	214	4	13.8	72 9	R48968	Gastrin-releasing pep	2.47e+03
142	5	17.2	931 15	R79009	Human DNA repair prot	2.62e+02	215	4	13.8	74 3	P61697	Sequence of enzyme do	2.47e+03
143	5	17.2	933 8	R41343	Bovine herpesvirus ty	2.62e+02	216	4	13.8	74 3	P61696	Sequence of enzyme do	2.47e+03
144	5	17.2	933 5	R27807	Bovine herpesvirus t	2.62e+02	217	4	13.8	74 3	P61701	Sequence of enzyme do	2.47e+03
145	5	17.2	993 31	W48720	Human mitochondrial i	2.62e+02	218	4	13.8	75 39	W88224	Human CD27 binding pr	2.47e+03
146	5	17.2	1022 3	P50097	Fusion protein of 59-	2.62e+02	219	4	13.8	75 39	W67948	Fragment of human sc	2.47e+03
147	5	17.2	1049 37	W73062	Rat GRP2.	2.62e+02	220	4	13.8	76 19	R84420	Hepatitis C virus EI	2.47e+03
148	5	17.2	1121 7	R36728	Ubiquitin fusion prot	2.62e+02	221	4	13.8	81 36	W80358	Sugi allergen protein	2.47e+03
149	5	17.2	1135 30	W33808	Mouse ataxin-2.	2.62e+02	222	4	13.8	83 39	W88234	Rana pipiens RNase A	2.47e+03
150	5	17.2	1144 13	R76059	Mycoplasma pirum adhe	2.62e+02	223	4	13.8	83 20	W01211	Serine protease PfSP1	2.47e+03
151	5	17.2	1163 2	R07120	P150.95 alpha subunit	2.62e+02	224	4	13.8	84 39	W87597	Guarmerin/Bufoin II f	2.47e+03
152	5	17.2	1163 32	W65091	Human Beta-integrin C	2.62e+02	225	4	13.8	84 36	W80650	S. pneumoniae transpo	2.47e+03
153	5	17.2	1186 32	W61030	Mus musculus Uty prot	2.62e+02	226	4	13.8	84 37	W74745	Human secreted protei	2.47e+03
154	5	17.2	1199 5	R26188	Hvra antigen.	2.62e+02	227	4	13.8	86 34	W62690	Streptococcus pneumo	2.47e+03
155	5	17.2	1217 10	R49832	Beat-galactosidase/he	2.62e+02	228	4	13.8	87 34	W29679	Homo sapiens clone DN	2.47e+03
156	5	17.2	1257 29	W41781	B. lactofermentum alp	2.62e+02	229	4	13.8	89 2	R11778	Enzyme donor polypept	2.47e+03
157	5	17.2	1257 17	R87628	Alpha-ketoglutaric ac	2.62e+02	230	4	13.8	90 34	W77647	Glutathione peroxidas	2.47e+03
158	5	17.2	1312 30	W33807	Human ataxin-2.	2.62e+02	231	4	13.8	92 39	W82717	Human Act-2 protein.	2.47e+03
159	5	17.2	1313 33	W60213	Spinoocerebellar ataxi	2.62e+02	232	4	13.8	92 35	W76225	Human chemokine MIP-1	2.47e+03
160	5	17.2	1334 3	R14481	LktA::lacZ fusion prod	2.62e+02	233	4	13.8	93 31	W27903	Amino acid sequence o	2.47e+03
161	5	17.2	1334 10	R50290	LktA::lacZ fusion pro	2.62e+02	234	4	13.8	97 37	W81663	M. tuberculosis immun	2.47e+03
162	5	17.2	1403 36	W79574	Leukotoxin 101.	2.62e+02	235	4	13.8	99 39	W73770	M. tuberculosis antig	2.47e+03

236	4	13.8	99 39	W73660	M. tuberculosis antig	2.47e+03	309	4	13.8	173 31	W28351	Staphylococcus aureus	2.47e+03
237	4	13.8	99 39	W73663	M. tuberculosis antig	2.47e+03	310	4	13.8	176 22	W24635	H. pylori cytoplasmic	2.47e+03
238	4	13.8	99 39	W73773	M. tuberculosis antig	2.47e+03	311	4	13.8	179 39	W81348	Melon expansin Cnx1	2.47e+03
239	4	13.8	100 30	W26793	Human factor VIII lig	2.47e+03	312	4	13.8	180 38	W89948	Antigen 3 from cluse	2.47e+03
240	4	13.8	100 7	R38597	Human lambda light ch	2.47e+03	313	4	13.8	181 39	W89980	Protein encoded by cl	2.47e+03
241	4	13.8	102 39	W95516	Fatty acid desaturase	2.47e+03	314	4	13.8	181 32	W56618	Human BY55 protein.	2.47e+03
242	4	13.8	102 38	W85126	A putative desaturase	2.47e+03	315	4	13.8	183 39	W89689	Murine osteogenic pro	2.47e+03
243	4	13.8	102 37	W84147	Phaeodactylum tricorn	2.47e+03	316	4	13.8	183 8	R44756	Murine osteogenic pro	2.47e+03
244	4	13.8	103 31	W58484	Human NEWM antibody 1	2.47e+03	317	4	13.8	183 35	W72006	HSV-2 strain SB5 Cont	2.47e+03
245	4	13.8	103 4	R24132	Human HLA SB-beta.	2.47e+03	318	4	13.8	186 34	W77774	Staphylococcus aureus	2.47e+03
246	4	13.8	103 7	R38605	NEWM light chain.	2.47e+03	319	4	13.8	187 38	W89878	Antigen 1 from cluse	2.47e+03
247	4	13.8	104 39	W88233	Rana pipiens RNase A	2.47e+03	320	4	13.8	187 38	W83903	Porcine interferon-be	2.47e+03
248	4	13.8	104 2	R12344	Protein with activity	2.47e+03	321	4	13.8	187 38	W73232	Porcine interferon-be	2.47e+03
249	4	13.8	105 7	R38600	Human heavy chain sub	2.47e+03	322	4	13.8	189 39	W88222	Human CD27 binding pr	2.47e+03
250	4	13.8	106 20	R99947	Mutated OCIF, OCIF-CC	2.47e+03	323	4	13.8	189 36	W80655	S. pneumoniae protein	2.47e+03
251	4	13.8	106 7	R38595	Human kappa light cha	2.47e+03	324	4	13.8	192 13	R69636	Hepatitis C virus env	2.47e+03
252	4	13.8	107 22	W07439	Anti-DNA antibody 9f1	2.47e+03	325	4	13.8	192 13	R69637	Hepatitis C virus env	2.47e+03
253	4	13.8	107 1	P93040	Recombinant chimeric	2.47e+03	326	4	13.8	192 39	W87805	Murine Bcl-2 associat	2.47e+03
254	4	13.8	109 16	R79858	Anti-EGFR antibody li	2.47e+03	327	4	13.8	192 39	W87808	Murine Bcl-2 associat	2.47e+03
255	4	13.8	111 9	R45606	Monoclonal antibody G	2.47e+03	328	4	13.8	192 18	R89544	Hepatitis C virus iso	2.47e+03
256	4	13.8	111 23	W13528	Anti-melanoma light c	2.47e+03	329	4	13.8	192 18	R89547	Hepatitis C virus iso	2.47e+03
257	4	13.8	111 17	R89506	Vllambda for antibody	2.47e+03	330	4	13.8	192 18	R89508	Hepatitis C virus iso	2.47e+03
258	4	13.8	111 31	W58485	Human KOL antibody li	2.47e+03	331	4	13.8	192 34	W64369	Mycobacterium tubercu	2.47e+03
259	4	13.8	111 2	R12263	Anti-human Rhd FOG-B	2.47e+03	332	4	13.8	192 37	W81734	M. tuberculosis immun	2.47e+03
260	4	13.8	111 23	W08491	C6 human sfv antibody	2.47e+03	333	4	13.8	192 13	R59639	Hepatitis C virus env	2.47e+03
261	4	13.8	112 23	W13523	Anti-melanoma light c	2.47e+03	334	4	13.8	193 12	R83366	Hepatitis C virus cor	2.47e+03
262	4	13.8	112 31	W08488	C6 human sfv antibody	2.47e+03	335	4	13.8	193 7	R33987	HC-J1 E1 protein.	2.47e+03
263	4	13.8	112 31	W60153	M. vaccae antigen GV-	2.47e+03	336	4	13.8	196 38	W30007	Expressed antigen for	2.47e+03
264	4	13.8	113 2	R68614	Anti-phenylloxazalone	2.47e+03	337	4	13.8	199 21	W09793	Natural killer cell e	2.47e+03
265	4	13.8	113 24	W25749	Human lctex-1.	2.47e+03	338	4	13.8	200 22	W20452	H. pylori secreted or	2.47e+03
266	4	13.8	114 2	R08431	HIV-1 Rev M7 mutant t	2.47e+03	339	4	13.8	201 21	W13274	Human von Willebrand	2.47e+03
267	4	13.8	115 1	R04134	Anti-Leu 3a light cha	2.47e+03	340	4	13.8	202 38	W88629	Secreted protein enco	2.47e+03
268	4	13.8	119 34	W64375	Mycobacterium tubercu	2.47e+03	341	4	13.8	202 39	W87822	Human secreted protei	2.47e+03
269	4	13.8	119 1	P81366	Light chain variable	2.47e+03	342	4	13.8	203 39	W67898	Human secreted protei	2.47e+03
270	4	13.8	121 3	P50651	Sequence encoded by b	2.47e+03	343	4	13.8	209 19	W04250	10A1 murine leukemia	2.47e+03
271	4	13.8	124 31	W88223	Human CD27 binding pr	2.47e+03	344	4	13.8	209 4	R20612	C10-14 NANBH-specific	2.47e+03
272	4	13.8	124 31	W28056	Staphylococcus aureus	2.47e+03	345	4	13.8	210 39	R86307	Kidney injury associa	2.47e+03
273	4	13.8	127 8	R39151	Staphylococcus aureus	2.47e+03	346	4	13.8	210 38	W89990	Expressed antigen for	2.47e+03
274	4	13.8	127 31	W46304	Humanised antibody 77	2.47e+03	347	4	13.8	210 38	W89867	Antigen 2 from cluse	2.47e+03
275	4	13.8	127 33	W56280	Babesia microti BMNI-	2.47e+03	348	4	13.8	210 22	W17902	Cellulase from Bacil	2.47e+03
276	4	13.8	128 31	W11773	Early onset Alzheimer	2.47e+03	349	4	13.8	211 38	W88548	Secreted protein enco	2.47e+03
277	4	13.8	128 32	W63829	Human Mab #117-10C li	2.47e+03	350	4	13.8	211 10	R53808	F19.9 V-min construct	2.47e+03
278	4	13.8	131 27	W19533	HSV 863 antibody ligh	2.47e+03	351	4	13.8	211 38	W67614	Hepatitis C virus E1	2.47e+03
279	4	13.8	131 1	P90543	Amino acids sequence	2.47e+03	352	4	13.8	211 17	R88556	Marrow stromal cell p	2.47e+03
280	4	13.8	132 13	R70189	Mouse Mab 3B9 light c	2.47e+03	353	4	13.8	214 24	W27091	Mouse monoclonal anti	2.47e+03
281	4	13.8	136 20	W03083	Staphylokinase deriva	2.47e+03	354	4	13.8	217 39	W82696	P. neopolydactyla typ	2.47e+03
282	4	13.8	136 20	W03087	Staphylokinase deriva	2.47e+03	355	4	13.8	217 37	W81518	Protein tyrosine phos	2.47e+03
283	4	13.8	136 28	W44695	Staphylokinase mutant	2.47e+03	356	4	13.8	217 34	W62677	Streptococcus pneumo	2.47e+03
284	4	13.8	136 20	W03093	Staphylokinase deriva	2.47e+03	357	4	13.8	220 9	R42340	Subunit of the human	2.47e+03
285	4	13.8	136 20	W03096	Staphylokinase deriva	2.47e+03	358	4	13.8	220 36	W1362	Human cytokine/steroi	2.47e+03
286	4	13.8	137 8	R39149	Staphylokinase SAK-Cp	2.47e+03	359	4	13.8	222 29	W40074	Human eosinophil gran	2.47e+03
287	4	13.8	137 8	R39153	Staphylokinase SAKM26	2.47e+03	360	4	13.8	225 39	W81346	Tomato expansin LeEx1	2.47e+03
288	4	13.8	137 8	R39150	Staphylokinase SAK-ST	2.47e+03	361	4	13.8	226 38	W73498	Von Willebrand factor	2.47e+03
289	4	13.8	140 39	W86331	Kidney injury associa	2.47e+03	362	4	13.8	226 35	W79397	Staphylococcus aureus	2.47e+03
290	4	13.8	140 9	R49659	Sequence of Heliothis	2.47e+03	363	4	13.8	230 39	W73517	KSHV protease.	2.47e+03
291	4	13.8	146 35	R12539	Cytotoxic T Lymphocyt	2.47e+03	364	4	13.8	230 17	R98144	Tne DNA polymerase C-	2.47e+03
292	4	13.8	150 31	W27740	Human acid sequence o	2.47e+03	365	4	13.8	230 17	R98145	Human secreted protei	2.47e+03
293	4	13.8	150 13	R77879	Human Cdn-1(60-211).	2.47e+03	366	4	13.8	235 37	W5214	Human secreted protei	2.47e+03
294	4	13.8	152 31	W00195	Mutant eukaryotic ini	2.47e+03	367	4	13.8	235 22	W20420	H. pylori cell envelo	2.47e+03
295	4	13.8	154 21	W00196	Mutant eukaryotic ini	2.47e+03	368	4	13.8	236 21	W01821	Primatised anti-human	2.47e+03
296	4	13.8	154 36	W80628	S. pneumoniae fimbria	2.47e+03	369	4	13.8	238 38	W81981	Ehrlichia sp. E82.3 p	2.47e+03
297	4	13.8	155 34	W55013	Novel haemopoietin re	2.47e+03	370	4	13.8	238 39	W83034	Anti-Fas humanised an	2.47e+03
298	4	13.8	155 33	W64716	Human TIABP1 protein.	2.47e+03	371	4	13.8	238 39	W83035	Anti-Fas humanised an	2.47e+03
299	4	13.8	158 34	W64718	Human TIABP1 genomic	2.47e+03	372	4	13.8	238 39	W83033	Anti-Fas humanised an	2.47e+03
300	4	13.8	158 34	W64718	Human TIABP1 genomic	2.47e+03	373	4	13.8	238 39	W83042	Anti-Fas MAB HFE7A li	2.47e+03
301	4	13.8	158 21	W20088	H. pylori protein, 10	2.47e+03	374	4	13.8	238 39	W83032	Anti-Fas humanised an	2.47e+03
302	4	13.8	158 15	R90679	Human SFT-SSX1 fusion	2.47e+03	375	4	13.8	238 39	W83031	Anti-Fas humanised an	2.47e+03
303	4	13.8	161 38	W82632	Ehrlichia sp. HGE-13	2.47e+03	376	4	13.8	238 4	R22053	Brassicin synthase fac	2.47e+03
304	4	13.8	163 2	R12137	S.aureus Staphylokina	2.47e+03	377	4	13.8	238 10	R54028	Synthase factor B (cl	2.47e+03
305	4	13.8	166 6	R30062	HCV E1 (envelope prot	2.47e+03	378	4	13.8	239 30	W40364	Rat Fas ligand-like p	2.47e+03
306	4	13.8	166 35	W59886	Amino acid sequence o	2.47e+03	379	4	13.8	240 9	R45391	Alcaligenes bronchise	2.47e+03
307	4	13.8	172 36	W72084	HSV-2 strain SB5 Cont	2.47e+03	380	4	13.8	240 1	R06827	Thrombomodulin analog	2.47e+03
308	4	13.8	173 38	W73428	Human secreted protei	2.47e+03	381	4	13.8	240 30	W40362	Human Fas ligand-like	2.47e+03

382	4	13.8	241	3	R13886	NGF with pro-region a	2.47e+03	455	4	13.8	312	33	W65038	Forsythia pinoresinol	2.47e+03
383	4	13.8	241	33	W48886	Human prepro-nerve gr	2.47e+03	456	4	13.8	317	38	W88758	Polypeptide fragment	2.47e+03
384	4	13.8	242	38	W90006	Expressed antigen for	2.47e+03	457	4	13.8	319	9	R45330	Anti-HCV antibody rea	2.47e+03
385	4	13.8	242	1	R06828	Thrombomodulin analog	2.47e+03	458	4	13.8	320	39	W89744	Staphylococcus aureus	2.47e+03
386	4	13.8	243	39	W89905	Antigen 1 from cluste	2.47e+03	459	4	13.8	320	6	R31624	Mature vwf subunit am	2.47e+03
387	4	13.8	245	38	W88654	Secreted protein enco	2.47e+03	460	4	13.8	320	35	W70220	Leishmania antigen lm	2.47e+03
388	4	13.8	247	26	W19339	Hybrid polypeptide p1	2.47e+03	461	4	13.8	320	35	W70236	Leishmania antigen pr	2.47e+03
389	4	13.8	247	37	W72384	Pathogen response pro	2.47e+03	462	4	13.8	324	39	W83390	Caenorhabditis elegans	2.47e+03
390	4	13.8	247	8	R39511	C terminal fragment o	2.47e+03	463	4	13.8	325	2	P70428	Polypeptide encoded b	2.47e+03
391	4	13.8	248	38	W73284	Human proteasome subu	2.47e+03	464	4	13.8	325	13	R75225	Human IgE Fc chain (a	2.47e+03
392	4	13.8	253	19	W05393	Human amyloid precurs	2.47e+03	465	4	13.8	326	20	R99840	Mutated OCIF, OCIF-DD	2.47e+03
393	4	13.8	253	21	W11945	p53 binding protein p	2.47e+03	466	4	13.8	327	31	W48626	Human adenomatous pol	2.47e+03
394	4	13.8	253	1	R04241	Thrombin-binding poly	2.47e+03	467	4	13.8	331	35	W76223	Human chemokine MIP-1	2.47e+03
395	4	13.8	254	30	W41987	Flea serine protease	2.47e+03	468	4	13.8	336	39	W86365	Human DNAX toll-like	2.47e+03
396	4	13.8	258	1	R06619	Modified Pseudomonas	2.47e+03	469	4	13.8	337	33	W61155	Haemophilus influenza	2.47e+03
397	4	13.8	258	1	R06686	Modified Pseudomonas	2.47e+03	470	4	13.8	337	14	R79217	pHCV351-encoded APP-H	2.47e+03
398	4	13.8	258	1	R06618	Modified Pseudomonas	2.47e+03	471	4	13.8	338	26	W36045	Tobacco cinnamoyl CoA	2.47e+03
399	4	13.8	258	23	W08487	C6 human sfv antibody	2.47e+03	472	4	13.8	338	26	W36043	Poplar cinnamoyl CoA	2.47e+03
400	4	13.8	260	34	W38609	Streptococcus pneumon	2.47e+03	473	4	13.8	343	39	W87636	The fusion protein GS	2.47e+03
401	4	13.8	261	38	W88759	Polypeptide fragment	2.47e+03	474	4	13.8	343	14	R73142	Coprinus cinereus per	2.47e+03
402	4	13.8	261	2	R11599	Beta-1,3-glucanase.	2.47e+03	475	4	13.8	343	14	R73145	Coprinus cinereus per	2.47e+03
403	4	13.8	261	37	W72933	Mycobacterium tubercu	2.47e+03	476	4	13.8	343	14	R73146	Coprinus cinereus per	2.47e+03
404	4	13.8	262	5	R28836	VP2 Thrlfile of HRV s	2.47e+03	477	4	13.8	343	14	R73154	Coprinus cinereus per	2.47e+03
405	4	13.8	264	17	R95570	Intracellular binding	2.47e+03	478	4	13.8	343	14	R73153	Coprinus cinereus per	2.47e+03
406	4	13.8	272	15	R80401	Anti-HIV-1 MAb 447 sc	2.47e+03	479	4	13.8	343	14	R73134	Coprinus cinereus per	2.47e+03
407	4	13.8	272	21	W20101	H. pylori transporter	2.47e+03	480	4	13.8	343	14	R73126	Coprinus cinereus per	2.47e+03
408	4	13.8	272	15	R80395	Anti-HIV-1 MAb 447 sc	2.47e+03	481	4	13.8	343	14	R73151	Coprinus cinereus per	2.47e+03
409	4	13.8	272	15	R80393	Anti-HIV-1 MAb 447 sc	2.47e+03	482	4	13.8	343	14	R73149	Coprinus cinereus per	2.47e+03
410	4	13.8	272	15	R80393	Anti-HIV-1 MAb 447 sc	2.47e+03	483	4	13.8	344	8	R40918	Bacillus subtilis ino	2.47e+03
411	4	13.8	272	15	R80389	Anti-HIV-1 MAb 447 sc	2.47e+03	484	4	13.8	345	39	W73643	Human placenta short	2.47e+03
412	4	13.8	272	15	R80388	Anti-HIV-1 MAb 447 sc	2.47e+03	485	4	13.8	346	21	W13841	Phage resistance gene	2.47e+03
413	4	13.8	272	15	R80387	Anti-HIV-1 MAb 447 sc	2.47e+03	486	4	13.8	349	39	W83928	Human FTHMA-070 parti	2.47e+03
414	4	13.8	273	23	W22110	P. tetragonolobus mat	2.47e+03	487	4	13.8	349	39	W54365	Human retinitis pigme	2.47e+03
415	4	13.8	275	8	R51088	Subtilisin BN <sup>+</sup> (P172	2.47e+03	488	4	13.8	357	32	W60227	Bacillus thuringiensis	2.47e+03
416	4	13.8	275	8	R51085	Subtilisin BN <sup>+</sup> mutei	2.47e+03	489	4	13.8	359	21	W14523	Human chimeric fucosy	2.47e+03
417	4	13.8	275	10	R51927	Bacillus subtilis sub	2.47e+03	490	4	13.8	359	21	W14524	Human chimeric fucosy	2.47e+03
418	4	13.8	275	2	R10210	Mutant subtilisin pol	2.47e+03	491	4	13.8	359	21	W14522	Human chimeric fucosy	2.47e+03
419	4	13.8	275	10	R52647	B. subtilis subtilisin	2.47e+03	492	4	13.8	359	20	R99939	Mutated OCIF, OCIF-DC	2.47e+03
420	4	13.8	275	2	R12140	Fibrinolytic protein.	2.47e+03	493	4	13.8	359	21	W14529	Human chimeric fucosy	2.47e+03
421	4	13.8	275	10	R52646	B. subtilis subtilisin	2.47e+03	494	4	13.8	360	32	W37920	Homo sapiens Dnae II	2.47e+03
422	4	13.8	275	10	R52645	B. subtilis subtilisin	2.47e+03	495	4	13.8	360	39	W06797	CCAAAT/Enhancer bindin	2.47e+03
423	4	13.8	275	2	R10204	Mutant subtilisin pol	2.47e+03	496	4	13.8	361	39	W73642	Rat adenosine kinase	2.47e+03
424	4	13.8	275	2	R10206	Mutant subtilisin pol	2.47e+03	497	4	13.8	361	3	R13749	GDP-Fuc:[beta-D-Gal(1	2.47e+03
425	4	13.8	275	1	P90095	Subtilisin variant	2.47e+03	498	4	13.8	362	39	W40817	Human bone morphogene	2.47e+03
426	4	13.8	275	3	R13467	Cc protein.	2.47e+03	499	4	13.8	362	39	W73644	Human placenta long f	2.47e+03
427	4	13.8	276	38	W81577	EDG-1-like G-protein	2.47e+03	500	4	13.8	362	37	W72963	Orange pectin methyl	2.47e+03
428	4	13.8	278	27	W31609	2,5-diketo-D-gluconic	2.47e+03	501	4	13.8	363	38	W88563	Secreted protein enco	2.47e+03
429	4	13.8	278	27	W31608	2,5-diketo-D-gluconic	2.47e+03	502	4	13.8	363	13	R75422	Coprinus peroxidase.	2.47e+03
430	4	13.8	278	26	W31625	2,5-diketo-D-gluconic	2.47e+03	503	4	13.8	363	29	W47408	Wild type guinea pig	2.47e+03
431	4	13.8	278	1	R05162	Sequence of human bon	2.47e+03	504	4	13.8	363	14	R78232	Coprinus cinereus per	2.47e+03
432	4	13.8	278	10	R49928	2,5-diketo-D-gluconic	2.47e+03	505	4	13.8	363	22	W14139	Arabidopsis thaliana	2.47e+03
433	4	13.8	278	26	W31621	2,5-diketo-D-gluconic	2.47e+03	506	4	13.8	364	4	R24151	Arthromyces ramosus p	2.47e+03
434	4	13.8	278	26	W31618	2,5-diketo-D-gluconic	2.47e+03	507	4	13.8	365	29	W47410	Wild type human L-aspi	2.47e+03
435	4	13.8	278	26	W31616	2,5-diketo-D-gluconic	2.47e+03	508	4	13.8	367	14	R99218	pHCV167-encoded prote	2.47e+03
436	4	13.8	278	1	P94624	2,5-diketo-D-gluconic	2.47e+03	509	4	13.8	370	38	W88777	Polypeptide fragment	2.47e+03
437	4	13.8	280	37	W13633	HIV POL/NEF epitopes.	2.47e+03	510	4	13.8	372	37	W36616	Celebus macaque Zcyto	2.47e+03
438	4	13.8	282	3	R12302	ORF-4 of Herpes Virus	2.47e+03	511	4	13.8	374	38	W73434	Human secreted protei	2.47e+03
439	4	13.8	286	14	R80827	Human MGFDF-3.	2.47e+03	512	4	13.8	381	39	W73505	Decay accelerating fa	2.47e+03
440	4	13.8	289	33	W5072	Streptococcus pneumon	2.47e+03	513	4	13.8	381	2	P71060	Wild-type subtilisin	2.47e+03
441	4	13.8	290	8	R41272	vWF fragment Arg441-A	2.47e+03	514	4	13.8	381	24	W26317	Human decay accelerat	2.47e+03
442	4	13.8	290	8	R41271	vWF fragment Arg441-A	2.47e+03	515	4	13.8	381	1	R03737	Subtilisin gene produ	2.47e+03
443	4	13.8	290	8	R41270	vWF fragment Arg441-A	2.47e+03	516	4	13.8	383	39	W85004	Amino acid sequence o	2.47e+03
444	4	13.8	293	35	W80647	S. pneumoniae transpo	2.47e+03	517	4	13.8	383	28	W26751	Bovine whey protein.	2.47e+03
445	4	13.8	293	3	P60680	Prepro- and mature in	2.47e+03	518	4	13.8	387	37	W76987	Human EGR-3 protein.	2.47e+03
446	4	13.8	300	32	W37928	Amino acid sequence o	2.47e+03	519	4	13.8	390	27	W34548	Anti-dorsalising morp	2.47e+03
447	4	13.8	307	4	P40036	Sequence encoded by t	2.47e+03	520	4	13.8	390	13	R69516	Prostaglandin-EP3-alp	2.47e+03
448	4	13.8	309	39	W82496	S. pneumoniae 37-kDa	2.47e+03	521	4	13.8	391	1	P94154	Beta-ketothiolase.	2.47e+03
449	4	13.8	310	39	W83324	Single chain Apo-2 an	2.47e+03	522	4	13.8	393	20	R99948	Mutated OCIF, OCIF-CB	2.47e+03
450	4	13.8	311	34	W64371	Mycobacterium tubercu	2.47e+03	523	4	13.8	393	32	W60133	M. vaccae potD homolo	2.47e+03
451	4	13.8	311	37	W81756	M. tuberculosis immun	2.47e+03	524	4	13.8	395	31	W57636	Modified TR1 receptor	2.47e+03
452	4	13.8	312	38	W73398	Human secreted protei	2.47e+03	525	4	13.8	396	3	R14241	Human BMP-2A encoded	2.47e+03
453	4	13.8	312	32	W65039	Forsythia pinoresinol	2.47e+03	526	4	13.8	396	39	W86680	Human osteogenic prot	2.47e+03
454	4	13.8	312	32	W65040	Forsythia pinoresinol	2.47e+03	527	4	13.8	396	6	R29713	P62, tyrosine Kinase	2.47e+03

528	4	13.8	396	8	R44747	Osteogenic protein CB	2.47e+03
529	4	13.8	396	35	W64771	Serine threonine kina	2.47e+03
530	4	13.8	398	11	R60817	Human type II IL-1 re	2.47e+03
531	4	13.8	400	39	W37364	M. tuberculosis antig	2.47e+03
532	4	13.8	400	39	W37364	M. tuberculosis antig	2.47e+03
533	4	13.8	401	39	W83926	Human FTHMA-070 prote	2.47e+03
534	4	13.8	402	38	W81966	Human P selectin liga	2.47e+03
535	4	13.8	402	38	W81966	Human P selectin liga	2.47e+03
536	4	13.8	403	3	R13990	P-selectin glycoprote	2.47e+03
537	4	13.8	404	15	R90675	Sialyltransferase mut	2.47e+03
538	4	13.8	406	38	W85089	Human chromosome-18 s	2.47e+03
539	4	13.8	406	10	R55707	Thyroid hormone recep	2.47e+03
540	4	13.8	407	34	W56458	Sialyltransferase	2.47e+03
541	4	13.8	409	13	R67552	Human growth differen	2.47e+03
542	4	13.8	410	17	R79227	Glucosylase	2.47e+03
543	4	13.8	410	17	R85866	WD-40 domain-contg. I	2.47e+03
544	4	13.8	412	38	W81967	Human P selectin liga	2.47e+03
545	4	13.8	412	38	W89246	Shingomonas sp. stra	2.47e+03
546	4	13.8	412	12	R63789	Aspergillus aculeatus	2.47e+03
547	4	13.8	416	38	W88745	Secreted protein enco	2.47e+03
548	4	13.8	417	14	R79228	pHCV424-encoded APP-H	2.47e+03
549	4	13.8	418	39	W73574	GST-Msp I methylase f	2.47e+03
550	4	13.8	418	39	W73574	GST-Msp I methylase f	2.47e+03
551	4	13.8	418	25	W26545	Mouse Huntington's di	2.47e+03
552	4	13.8	420	22	W20972	H. pylori cytoplasmic	2.47e+03
553	4	13.8	425	17	R85887	WD-40 domain-contg. h	2.47e+03
554	4	13.8	425	19	W06798	Murine p154.	2.47e+03
555	4	13.8	425	22	W21013	H. pylori cytoplasmic	2.47e+03
556	4	13.8	427	10	R54855	EPSP synthase from E.	2.47e+03
557	4	13.8	427	4	R23067	Modified 5-enolpyruv	2.47e+03
558	4	13.8	427	1	R80867	Mutant araA gene enco	2.47e+03
559	4	13.8	427	5	R27794	EPSP synthase wild ty	2.47e+03
560	4	13.8	428	8	R22950	Human IgE heavy chain	2.47e+03
561	4	13.8	429	18	W03714	Human alpha-1C3 adren	2.47e+03
562	4	13.8	430	39	W89796	Staphylococcus aureus	2.47e+03
563	4	13.8	431	8	R38750	Sec Y protein.	2.47e+03
564	4	13.8	433	21	W11879	Hantavirus nuclear pr	2.47e+03
565	4	13.8	433	21	W11875	Hantavirus nuclear pr	2.47e+03
566	4	13.8	433	20	W09384	S-sugar protein from	2.47e+03
567	4	13.8	433	21	W11877	Hantavirus nuclear pr	2.47e+03
568	4	13.8	439	17	R85880	WD-40 domain-contg. Y	2.47e+03
569	4	13.8	443	39	W83372	Streptococcus pneumon	2.47e+03
570	4	13.8	444	14	R79372	Rhizobium DMP.	2.47e+03
571	4	13.8	445	17	R85885	WD-40 domain-contg. C	2.47e+03
572	4	13.8	445	34	W68150	Diaminopimelate decar	2.47e+03
573	4	13.8	446	21	W20095	H. pylori cytoplasmic	2.47e+03
574	4	13.8	448	12	R58668	Human PACAP receptor	2.47e+03
575	4	13.8	453	14	R79225	pHCV421-encoded APP-H	2.47e+03
576	4	13.8	453	34	W62857	Agrocyste pectinases phy	2.47e+03
577	4	13.8	454	17	R97870	Human Yes-associated	2.47e+03
578	4	13.8	455	5	R28757	Hepatocyte nuclear fa	2.47e+03
579	4	13.8	459	8	R42848	VP receptor protein.	2.47e+03
580	4	13.8	462	39	W73623	Human secreted protei	2.47e+03
581	4	13.8	462	9	R45350	Thrombomodulin analog	2.47e+03
582	4	13.8	462	17	R89383	Human semenogelin I.	2.47e+03
583	4	13.8	462	9	R45346	Thrombomodulin analog	2.47e+03
584	4	13.8	462	9	R45346	Thrombomodulin analog	2.47e+03
585	4	13.8	464	37	W72907	Mycobacterium tubercu	2.47e+03
586	4	13.8	466	38	W82398	Mouse G3BP protein.	2.47e+03
587	4	13.8	466	18	R90037	Human alpha-1C adren	2.47e+03
588	4	13.8	466	14	R73604	Human alpha 1C adren	2.47e+03
589	4	13.8	466	10	R52832	Sequence of human alp	2.47e+03
590	4	13.8	466	14	R58594	Alpha-1C adrenegic r	2.47e+03
591	4	13.8	466	12	R58679	Human alpha-1C adren	2.47e+03
592	4	13.8	466	18	R90038	Human alpha-1C adren	2.47e+03
593	4	13.8	467	38	W73400	Human secreted protei	2.47e+03
594	4	13.8	469	10	R54024	Synthase factor B.	2.47e+03
595	4	13.8	471	24	W26311	Human STCH chaperone	2.47e+03
596	4	13.8	472	39	W86327	Kidney injury associa	2.47e+03
597	4	13.8	475	8	R390750	Delta endotoxin.	2.47e+03
598	4	13.8	475	3	R20069	B.thuringiensis toxin	2.47e+03
599	4	13.8	475	14	R76114	Strain F552A1 acaride	2.47e+03
600	4	13.8	476	14	R78725	Mature thrombomodulin	2.47e+03
601	4	13.8	477	39	W87592	Rhodococcus strain EA	2.47e+03
602	4	13.8	477	25	W25153	Nsp7524111 restrictio	2.47e+03
603	4	13.8	477	12	R66067	Phosphorylated p66 db	2.47e+03
604	4	13.8	480	38	W74814	Human secreted protei	2.47e+03
605	4	13.8	481	9	R47579	Soluble Flk-2.	2.47e+03
606	4	13.8	483	30	W37064	HIV-1 breakthrough is	2.47e+03
607	4	13.8	483	4	P40694	Sequence encoded by p	2.47e+03
608	4	13.8	484	23	W22169	S.thermophilus expopol	2.47e+03
609	4	13.8	485	38	W85081	Esterase E008 from so	2.47e+03
610	4	13.8	487	28	W42979	Amyloid precursor pro	2.47e+03
611	4	13.8	487	28	W42978	Amyloid precursor pro	2.47e+03
612	4	13.8	492	1	P94263	Bovine adrenal gland	2.47e+03
613	4	13.8	493	4	P40065	Sequence of human imm	2.47e+03
614	4	13.8	494	6	R31888	Defective tyrosine ki	2.47e+03
615	4	13.8	494	14	R78726	Thrombomodulin with w	2.47e+03
616	4	13.8	496	11	R57349	Human cartilage matri	2.47e+03
617	4	13.8	496	31	W55962	Streptococcus pneumon	2.47e+03
618	4	13.8	498	15	R82242	Capsanthin-capsorubin	2.47e+03
619	4	13.8	499	2	P70056	Partial sequence of l	2.47e+03
620	4	13.8	500	2	R12772	CI inhibitor muten -	2.47e+03
621	4	13.8	500	23	W18214	Recombinant CI inhibi	2.47e+03
622	4	13.8	501	38	W85075	Amino acid sequence o	2.47e+03
623	4	13.8	501	38	W85088	Esterase E027.	2.47e+03
624	4	13.8	501	38	W85084	Esterase E015 from so	2.47e+03
625	4	13.8	501	38	W85083	Esterase E013 from so	2.47e+03
626	4	13.8	501	38	W85074	Amino acid sequence o	2.47e+03
627	4	13.8	501	38	W85079	Esterase E004 from so	2.47e+03
628	4	13.8	501	38	W85082	Esterase E010 from so	2.47e+03
629	4	13.8	501	17	R81471	Nocardia corallina al	2.47e+03
630	4	13.8	501	13	R65264	Rat G protein coupled	2.47e+03
631	4	13.8	503	10	R53366	Human Activin recepto	2.47e+03
632	4	13.8	508	26	W25747	Arabidopsis protoporp	2.47e+03
633	4	13.8	509	39	W95196	Bcl-2 interaction pro	2.47e+03
634	4	13.8	509	17	R97645	Mouse Sox-9 protein.	2.47e+03
635	4	13.8	509	37	W80955	Amino acid sequence o	2.47e+03
636	4	13.8	513	39	W87798	Aminopeptidase Slp e	2.47e+03
637	4	13.8	514	17	R85881	WD-40 domain-contg. Y	2.47e+03
638	4	13.8	515	37	W77106	Rat alpha18-adrenegri	2.47e+03
639	4	13.8	515	23	R15762	PAP-QM-CSF immunostim	2.47e+03
640	4	13.8	516	3	R14279	Sequence encoded by h	2.47e+03
641	4	13.8	516	36	W70356	Chlamydomonas protein	2.47e+03
642	4	13.8	516	4	P40135	Sequence of protein A	2.47e+03
643	4	13.8	518	5	R27644	Human calcium channel	2.47e+03
644	4	13.8	519	39	W89782	Staphylococcus aureus	2.47e+03
645	4	13.8	522	36	W72660	Canine herpes virus p	2.47e+03
646	4	13.8	523	11	R71976	Pertussis A.	2.47e+03
647	4	13.8	524	30	W46747	Arabidopsis thaliana	2.47e+03
648	4	13.8	528	35	W59882	Amino acid sequence o	2.47e+03
649	4	13.8	529	38	W82586	Human Patched-2 prote	2.47e+03
650	4	13.8	529	39	W73473	Human cancer-related	2.47e+03
651	4	13.8	530	33	W57877	Truncated transferin	2.47e+03
652	4	13.8	530	33	W57877	C. felis esterase pro	2.47e+03
653	4	13.8	530	33	W57855	C. felis esterase. nF	2.47e+03
654	4	13.8	533	32	W56484	Zea mays waxy gene gl	2.47e+03
655	4	13.8	535	32	W60243	Amino acid sequence o	2.47e+03
656	4	13.8	536	37	W81805	Porcine ZPB protein.	2.47e+03
657	4	13.8	539	31	W48358	Fractonated X irradi	2.47e+03
658	4	13.8	540	20	W07871	GPB (or Grb14), a sig	2.47e+03
659	4	13.8	543	30	W52186	POL region of p14gaq	2.47e+03
660	4	13.8	544	31	W51256	Maize proto-porphyrin	2.47e+03
661	4	13.8	544	38	W85028	Cdk2-green fluoresce	2.47e+03
662	4	13.8	544	26	W25748	Maize protoporphyrino	2.47e+03
663	4	13.8	547	28	W43013	Truncated transferin	2.47e+03
664	4	13.8	548	39	W89189	Alternatively spliced	2.47e+03
665	4	13.8	549	38	W70898	Acetyl-coenzyme A tra	2.47e+03
666	4	13.8	550	39	W81351	Human guanine nucleot	2.47e+03
667	4	13.8	551	19	R97422	Fis1 gene product.	2.47e+03
668	4	13.8	551	13	R71376	Human cystathionine b	2.47e+03
669	4	13.8	551	14	R83663	Human p58 kinase mRNA	2.47e+03
670	4	13.8	553	12	R58660	Human PACAP receptor	2.47e+03
671	4	13.8	553	37	W81996	A. thaliana trehalase	2.47e+03
672	4	13.8	559	37	W85862	Pseudomonas species p	2.47e+03
673	4	13.8	566	39	W87531	An antitumour protein	2.47e+03





820	4	13.8	893 38	W83983	Tne DNA polymerase mu	2.47e+03	893	4	13.8	1254 24	W24575	Merozoite apical-end-1	2.47e+03
821	4	13.8	893 38	W83993	Tne DNA polymerase mu	2.47e+03	894	4	13.8	1254 2	R07503	Merozoite apical-end-1	2.47e+03
822	4	13.8	893 38	W83980	Tne DNA polymerase mu	2.47e+03	895	4	13.8	1274 39	W89253	Human ALP	2.47e+03
823	4	13.8	893 38	W83981	Tne DNA polymerase mu	2.47e+03	896	4	13.8	1275 39	W82594	H. contortus pEP-A pr	2.47e+03
824	4	13.8	893 38	W83987	Tne DNA polymerase mu	2.47e+03	897	4	13.8	1280 8	R44297	Sequence encoded by h	2.47e+03
825	4	13.8	893 38	W83986	Tne DNA polymerase mu	2.47e+03	898	4	13.8	1280 33	W48999	C-terminal single mut	2.47e+03
826	4	13.8	893 38	W83979	Tne DNA polymerase mu	2.47e+03	899	4	13.8	1280 33	W48997	Wild-type human P gly	2.47e+03
827	4	13.8	893 38	W83988	Tne DNA polymerase mu	2.47e+03	900	4	13.8	1280 2	P70452	Sequence encoded by h	2.47e+03
828	4	13.8	893 38	W83997	Tne DNA polymerase mu	2.47e+03	901	4	13.8	1288 24	P70452	Mouse alpha-1 collage	2.47e+03
829	4	13.8	893 38	W83998	Tne DNA polymerase mu	2.47e+03	902	4	13.8	1292 1	R05531	High density lipoprot	2.47e+03
830	4	13.8	893 38	W83996	Tne DNA polymerase mu	2.47e+03	903	4	13.8	1294 36	W75798	Arabidopsis ethylene	2.47e+03
831	4	13.8	893 38	W83976	Tne DNA polymerase mu	2.47e+03	904	4	13.8	1315 37	W75734	Human mdia Rho target	2.47e+03
832	4	13.8	893 38	W84004	Tne DNA polymerase mu	2.47e+03	905	4	13.8	1316 38	W87511	Human N-methyl-D-aspa	2.47e+03
833	4	13.8	893 38	W84009	Tne DNA polymerase mu	2.47e+03	906	4	13.8	1336 12	R66041	Human N-methyl-D-aspa	2.47e+03
834	4	13.8	893 38	W83991	Tne DNA polymerase mu	2.47e+03	907	4	13.8	1338 37	W80600	Helicobacter pylori C	2.47e+03
835	4	13.8	893 38	W83990	Tne DNA polymerase mu	2.47e+03	908	4	13.8	1342 36	W69406	Erbb-3 glycoprotein c	2.47e+03
836	4	13.8	893 38	W84007	Tne DNA polymerase mu	2.47e+03	909	4	13.8	1349 33	W59359	Human retinal degener	2.47e+03
837	4	13.8	893 38	W83994	Tne DNA polymerase mu	2.47e+03	910	4	13.8	1367 22	W19875	Murine flk-1 receptor	2.47e+03
838	4	13.8	893 38	W83995	Tne DNA polymerase mu	2.47e+03	911	4	13.8	1367 10	R54046	Sequence of murine fo	2.47e+03
839	4	13.8	893 38	W83985	Tne DNA polymerase mu	2.47e+03	912	4	13.8	1370 3	P60005	Sequence encoded by h	2.47e+03
840	4	13.8	893 38	W83988	Tne DNA polymerase mu	2.47e+03	913	4	13.8	1399 8	R38698	S-PRV-055 TGE virus g	2.47e+03
841	4	13.8	893 38	W83989	Tne DNA polymerase mu	2.47e+03	914	4	13.8	1404 7	R38304	Sequence of a serrate	2.47e+03
842	4	13.8	893 38	W84011	Tne DNA polymerase mu	2.47e+03	915	4	13.8	1417 17	R96955	Multi-drug resistance	2.47e+03
843	4	13.8	893 38	W83992	Tne DNA polymerase mu	2.47e+03	916	4	13.8	1471 27	W23414	Human B-domain delete	2.47e+03
844	4	13.8	893 38	W84006	Tne DNA polymerase mu	2.47e+03	917	4	13.8	1472 17	R96954	Multi-drug resistance	2.47e+03
845	4	13.8	893 38	W84008	Tne DNA polymerase mu	2.47e+03	918	4	13.8	1477 12	R63506	Haemophilus high mole	2.47e+03
846	4	13.8	893 38	W84005	Tne DNA polymerase mu	2.47e+03	919	4	13.8	1482 17	R96953	Multi-drug resistance	2.47e+03
847	4	13.8	896 12	R63533	Human HP-1376 cell-de	2.47e+03	920	4	13.8	1489 17	R96952	Multi-drug resistance	2.47e+03
848	4	13.8	899 17	R90987	Nitrate reductase	2.47e+03	921	4	13.8	1501 14	R72858	Rat receptor type-pro	2.47e+03
849	4	13.8	899 37	W44886	Chimeric polypurine-	2.47e+03	922	4	13.8	1516 1	P80255	Modified factor VIII:	2.47e+03
850	4	13.8	901 2	R11994	Glutamate receptor 6	2.47e+03	923	4	13.8	1531 10	R54928	Multi-drug resistance	2.47e+03
851	4	13.8	904 17	R87147	Protocadherin clone 4	2.47e+03	924	4	13.8	1536 12	R63505	Haemophilus high mole	2.47e+03
852	4	13.8	907 37	W49032	Human eif4G-like prot	2.47e+03	925	4	13.8	1536 27	P30293	Non-typeable Haemophi	2.47e+03
853	4	13.8	922 5	R25578	Bordetella parapertus	2.47e+03	926	4	13.8	1541 1	R07304	IgA1 protease	2.47e+03
854	4	13.8	933 38	W06651	Phosphoenolpyruvic ac	2.47e+03	927	4	13.8	1571 32	W42087	Human Down syndrome-c	2.47e+03
855	4	13.8	935 3	R15050	Hybrid human insulin-	2.47e+03	928	4	13.8	1598 27	W30291	Non-typeable Haemophi	2.47e+03
856	4	13.8	936 20	W05719	Rhodospirillum rubrum sp.	2.47e+03	929	4	13.8	1634 9	R42452	Enzyme involved in ei	2.47e+03
857	4	13.8	943 28	W43039	Human secreted protein	2.47e+03	930	4	13.8	1664 1	P80264	Sequence of the Japan	2.47e+03
858	4	13.8	945 39	W73624	Human p50 protein DNA	2.47e+03	931	4	13.8	1684 3	R14948	Bacterial amylose A-1	2.47e+03
859	4	13.8	955 8	R42235	Human p50 protein DNA	2.47e+03	932	4	13.8	1704 9	R49657	Sequence of Heliothrix	2.47e+03
860	4	13.8	955 8	R42242	Human p50 protein DNA	2.47e+03	933	4	13.8	1717 27	W23331	Neuroblastoma indicat	2.47e+03
861	4	13.8	955 8	R42247	Human p50 protein DNA	2.47e+03	934	4	13.8	1757 39	W84351	Murine ubiquitin-prot	2.47e+03
862	4	13.8	955 8	R42248	Human p50 protein DNA	2.47e+03	935	4	13.8	1833 14	R79478	Mouse LfBP-2	2.47e+03
863	4	13.8	961 39	W83387	Caenorhabditis elegans	2.47e+03	936	4	13.8	1911 24	W27225	Human protein tyrosin	2.47e+03
864	4	13.8	962 31	R58701	Aminopeptidase H110D	2.47e+03	937	4	13.8	1972 20	W00024	Smooth muscle myosin	2.47e+03
865	4	13.8	964 18	R98371	Spinach debanching e	2.47e+03	938	4	13.8	2050 38	W73499	Von Willebrand factor	2.47e+03
866	4	13.8	967 14	R79222	pHCV141-encoded seque	2.47e+03	939	4	13.8	2098 16	R65863	Factor-VIII	2.47e+03
867	4	13.8	977 8	R51282	Helminth aminopeptida	2.47e+03	940	4	13.8	2237 33	W63142	Human calcium channel	2.47e+03
868	4	13.8	990 39	W90075	P. vulgaris chondroit	2.47e+03	941	4	13.8	2237 6	R33550	Sequence of the alpha	2.47e+03
869	4	13.8	990 30	W37786	Amino acid sequence o	2.47e+03	942	4	13.8	2251 14	R71009	Human neuronal calciu	2.47e+03
870	4	13.8	992 9	R44994	Murine flk-2 receptor	2.47e+03	943	4	13.8	2262 31	W56737	Calcium ion channel a	2.47e+03
871	4	13.8	993 9	R44995	Human flk-2 receptor	2.47e+03	944	4	13.8	2270 14	R71010	Human neuronal calciu	2.47e+03
872	4	13.8	993 12	R67816	Flk2 receptor protein	2.47e+03	945	4	13.8	2308 11	R57902	Human RPTP-beta	2.47e+03
873	4	13.8	995 4	R22234	Sequence encoded by I	2.47e+03	946	4	13.8	2332 2	P71729	Factor VIII:c variant	2.47e+03
874	4	13.8	1002 1	P81861	Sequence encoded by L	2.47e+03	947	4	13.8	2332 30	W53483	Human factor VIII:	2.47e+03
875	4	13.8	1004 32	W37856	Human polyhomocytic 1	2.47e+03	948	4	13.8	2332 28	W33225	Procoagulant-active h	2.47e+03
876	4	13.8	1006 2	R07115	Gene product of yeast	2.47e+03	949	4	13.8	2332 2	P71728	Factor VIII:c variant	2.47e+03
877	4	13.8	1013 20	W09398	Proteus vulgaris chon	2.47e+03	950	4	13.8	2337 32	W37878	Human calcium channel	2.47e+03
878	4	13.8	1014 35	W68474	HIV-1 strain YBF30 po	2.47e+03	951	4	13.8	2339 33	W63141	Human calcium channel	2.47e+03
879	4	13.8	1016 2	R08054	HIV-1 pol protein of	2.47e+03	952	4	13.8	2339 6	R33549	Sequence of the alpha	2.47e+03
880	4	13.8	1037 39	W67643	A serine/threonine pr	2.47e+03	953	4	13.8	2343 24	W11335	Active Factor VIII:C	2.47e+03
881	4	13.8	1038 38	W86246	Human BMP receptor ki	2.47e+03	954	4	13.8	2344 24	W11331	Active Factor VIII:C	2.47e+03
882	4	13.8	1038 38	W86246	Mouse BMP receptor ki	2.47e+03	955	4	13.8	2345 24	W11415	Active Factor VIII:C	2.47e+03
883	4	13.8	1039 38	W85014	Stat5-green fluoresce	2.47e+03	956	4	13.8	2346 24	W11351	Active Factor VIII:C	2.47e+03
884	4	13.8	1045 38	W85050	Stat5-green fluoresce	2.47e+03	957	4	13.8	2346 24	W11360	Active Factor VIII:C	2.47e+03
885	4	13.8	1055 22	W14319	Thermotoga neopolitan	2.47e+03	958	4	13.8	2347 24	W11467	Active Factor VIII:C	2.47e+03
886	4	13.8	1089 38	W82396	Human uhp protein #2	2.47e+03	959	4	13.8	2348 24	W11341	Active Factor VIII:C	2.47e+03
887	4	13.8	1106 1	P90127	Platelet derived grow	2.47e+03	960	4	13.8	2348 24	W11395	Active Factor VIII:C	2.47e+03
888	4	13.8	1168 23	W16326	Nematocidal toxin 167	2.47e+03	961	4	13.8	2348 24	W11449	Active Factor VIII:C	2.47e+03
889	4	13.8	1169 18	R89469	Collagen/BMP-2B fusio	2.47e+03	962	4	13.8	2349 24	W11401	Active Factor VIII:C	2.47e+03
890	4	13.8	1172 31	W37870	Human protein compri	2.47e+03	963	4	13.8	2349 24	W11460	Active Factor VIII:C	2.47e+03
891	4	13.8	1203 8	R41668	Human endothelial cel	2.47e+03	964	4	13.8	2350 24	W11359	Active Factor VIII:C	2.47e+03
892	4	13.8	1244 10	R54781	Leukotoxin ApIIIA	2.47e+03	965	4	13.8	2351 24	W11396	Active Factor VIII:C	2.47e+03

966 4 13.8 2351 24 W11427 Active Factor VIII:C 2.47e+03  
 967 4 13.8 2351 24 W11332 Active Factor VIII:C 2.47e+03  
 968 4 13.8 2351 24 W11347 Active Factor VIII:C 2.47e+03  
 969 4 13.8 2351 24 W10591 Factor VIII:C (Phe652 2.47e+03  
 970 4 13.8 2352 24 W11382 Active Factor VIII:C 2.47e+03  
 971 4 13.8 2352 24 W11426 Active Factor VIII:C 2.47e+03  
 972 4 13.8 2352 24 W11350 Active Factor VIII:C 2.47e+03  
 973 4 13.8 2352 24 W11459 Active Factor VIII:C 2.47e+03  
 974 4 13.8 2352 24 W11333 Active Factor VIII:C 2.47e+03  
 975 4 13.8 2352 24 W11334 Active Factor VIII:C 2.47e+03  
 976 4 13.8 2352 24 W11354 Active Factor VIII:C 2.47e+03  
 977 4 13.8 2352 24 W11337 Active Factor VIII:C 2.47e+03  
 978 4 13.8 2356 35 W68388 Clostridium difficile 2.47e+03  
 979 4 13.8 2386 21 W13153 S. pombe Rad3 polyep 2.47e+03  
 980 4 13.8 2409 3 R12609 Versican 2.47e+03  
 981 4 13.8 2547 37 W81511 Drosophila fat facets 2.47e+03  
 982 4 13.8 2555 37 W81512 Drosophila fat facets 2.47e+03  
 983 4 13.8 2625 30 W59887 Human telomerase 2.47e+03  
 984 4 13.8 2627 33 W61347 Human telomerase RNA 2.47e+03  
 985 4 13.8 2629 33 W61348 Mouse telomerase RNA 2.47e+03  
 986 4 13.8 2639 39 W73476 Grapevine leafroll vi 2.47e+03  
 987 4 13.8 2756 30 W37050 S. putrefaciens EPO b 2.47e+03  
 988 4 13.8 2894 5 R24440 Composite HCV HC-J1/C 2.47e+03  
 989 4 13.8 2955 2 R08124 Hepatitis C virus put 2.47e+03  
 990 4 13.8 3210 1 P81770 Deduced sequence enco 2.47e+03  
 991 4 13.8 3396 8 R43662 DEN1-S275/90 (ECACC V 2.47e+03  
 992 4 13.8 3433 23 W22017 Utrophin 2.47e+03  
 993 4 13.8 3685 3 P90373 Sequence encoded by h 2.47e+03  
 994 4 13.8 3898 2 R10473 Hog cholera virus gen 2.47e+03  
 995 4 13.8 4302 29 W33396 Human PKD1 polypeptid 2.47e+03  
 996 4 13.8 4302 19 W00870 Polycystic kidney dis 2.47e+03  
 997 4 13.8 4302 28 W23830 Human PKD1 protein 2.47e+03  
 998 4 13.8 4303 17 R90302 Polycystic kidney dis 2.47e+03  
 999 4 13.8 4866 37 W77410 Human ryanodin recept 2.47e+03  
 1000 4 13.8 5035 5 R25450 MH mutant porcine rya 2.47e+03

## ALIGNMENTS

RESULT 1  
 ID W04842 standard; peptide; 15 AA.  
 AC W04842;  
 DT 18-FEB-1997 (first entry)  
 DE Self epitope of desmoglein 3, implicated in autoimmune disease.  
 KW Tolerisation; self-epitope; antigen; autoimmune disease;  
 KW autoantigen; HLA; human leukocyte antigen; T-cell; thymocyte;  
 KW pemphigus vulgaris; desmoglein; multiple sclerosis;  
 KW herpes simplex virus; adenovirus; phosphononmutase;  
 KW human papillomavirus; Epstein-Barr virus; DNA polymerase;  
 KW influenza; haemagglutinin; reovirus; sigma protein.  
 OS Homo sapiens.  
 PN W09627387-A1.  
 PD 12-SEP-1996.  
 PF 07-MAR-1996; U03182.  
 PR 07-MAR-1995; US-400796.  
 PA (HARD ) HARVARD COLLEGE.  
 PI Strominger JL, Wucherpfennig KW;  
 DR WPI: 96-425218/42.  
 PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self  
 PT antigens - useful in disease treatment, and method for  
 PT identification of other self and non-self antigens implicated in  
 PT auto-immune disease  
 PS Claim 1: Page 38: 58pp; English.  
 CC Pharmaceutical preparations for tolerisation to antigens comprise  
 CC either an isolated human non-collagen or non-mysin basic protein  
 CC (MBP) polypeptide which is capable of tolerising an individual to an  
 CC autoantigen; or an isolated human pathogen polypeptide capable of  
 CC tolerising an individual to that polypeptide. In both cases, the  
 CC polypeptide (whether self or non-self) includes an amino acid  
 CC sequence corresponding to a sequence motif for a MHC class II  
 CC protein, such as HLA-DR, which is associated with a human autoimmune  
 CC disease and which binds to the polypeptide to activate autoreactive  
 CC T-cells in individuals with the autoimmune disease. This peptide is

CC derived from the human desmoglein 3 protein (amino acids 97-111)  
 CC and is implicated as a self epitope in pemphigus vulgaris. Peptides  
 CC derived from the human desmoglein protein are described in W04841-47.  
 SQ Sequence 15 AA;  
 Query Match 24.1%; Score 7; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 4 fvvdktnt 10  
 Qy 22 FVVDKNT 28  
 RESULT 2  
 ID R93961 standard; peptide; 19 AA.  
 AC R93961;  
 DT 03-JUL-1996 (first entry)  
 DE Peptide combining with anti-interepidermal cellular antibody.  
 KW anti-interepidermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 1295556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 96-045392/05.  
 PT Anti-interepidermal cellular antibody-combining peptide - which can  
 PT be immobilised on column to form adsorbent useful for treating  
 PT diseases related to the antibody  
 PS Example 1; Page 4; 7pp; Japanese.  
 CC New peptides are disclosed which contain at least 5 contiguous amino  
 CC acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe  
 CC Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg  
 CC Glu Glu (see R93960), the peptide not containing more than 50 residues.  
 CC The peptide combines with anti-interepidermal cellular antibody. It can  
 CC be immobilised on a carrier to prepare an adsorbent useful for the  
 CC treatment of diseases related to anti-interepidermal cellular antibody.  
 CC The present sequence is a specific example of the new peptides.  
 SQ Sequence 19 AA;  
 Query Match 24.1%; Score 7; DB 16; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 11 fvvdktnt 17  
 Qy 22 FVVDKNT 28  
 RESULT 3  
 ID R93962 standard; peptide; 21 AA.  
 AC R93962;  
 DT 03-JUL-1996 (first entry)  
 DE Peptide combining with anti-interepidermal cellular antibody.  
 KW anti-interepidermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 1295556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 96-045392/05.  
 PT Anti-interepidermal cellular antibody-combining peptide - which can  
 PT be immobilised on column to form adsorbent useful for treating  
 PT diseases related to the antibody  
 PS Example 2; Page 5; 7pp; Japanese.  
 CC New peptides are disclosed which contain at least 5 contiguous amino  
 CC acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe  
 CC Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg  
 CC Glu Glu (see R93960), the peptide not containing more than 50 residues.  
 CC The peptide combines with anti-interepidermal cellular antibody. It can  
 CC be immobilised on a carrier to prepare an adsorbent useful for the

CC treatment of diseases related to anti-interepidermal cellular antibody.  
 CC The present sequence is a specific example of the new peptides.,  
 SQ Sequence 21 AA;

Query Match 24.1%; Score 7; DB 16; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 fvdvknk 19  
 |||||  
 QY 22 FVVDKNT 28

RESULT 4  
 ID R93960 standard; peptide; 30 AA.  
 AC R93960;  
 DT 05-JUL-1996 (first entry)  
 DE Peptide combining with anti-interepidermal cellular antibody.  
 KW anti-interepidermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 129556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS) KURARAY CO LTD.  
 DR WPI: 96-045392/05.  
 PT Anti-interepidermal cellular antibody-combining peptide - which can be immobilised on column to form adsorbent useful for treating diseases related to the antibody  
 PS Claim 1; Page 2; 7pp; Japanese.  
 CC A new peptide is disclosed which contains at least 5 contiguous amino acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg Glu Glu (the present sequence), the peptide not containing more than 50 residues. The peptide combines with anti-interepidermal cellular antibody. It can be immobilised on a carrier to prepare an absorbent useful for the treatment of diseases related to anti-interepidermal cellular antibody.  
 SQ Sequence 30 AA;

Query Match 24.1%; Score 7; DB 16; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 fvdvknk 17  
 |||||  
 QY 22 FVVDKNT 28

RESULT 5  
 ID W07908 standard; protein; 614 AA.  
 AC W07908;  
 DT 29-JAN-1997 (first entry)  
 DE Pemphigus vulgaris antigen protein extracellular region.  
 KW Autoantibody; immunoglobulin G; IgG1; fusion protein; diagnosis; treatment; pemphigus vulgaris; PV; bulla; blister; skin disease; dermatology.  
 KW Homo sapiens.  
 OS J08188540-A.  
 PN 23-JUL-1996.  
 PF 30-JUN-1995; 165632.  
 PR 30-JUN-1994; JP-173291.  
 PA (NISHU) NISHUKAWA T.  
 DR WPI: 96-388562/39.  
 PT Fused protein recognised by pemphigus vulgaris autoantibody - useful to treat and diagnose pemphigus vulgaris  
 PS Claim 1; Page 7-9; 9pp; Japanese.  
 CC W07908 represents the human pemphigus vulgaris (PV) antigen extracellular region. The PV antigen is produced in patients with pemphigus vulgaris resulting in autoimmune disease. PV is a rare relapsing disease causing suprabasal, intra-epidermal bullae (vesicles) of the skin and mucous membranes, which is fatal if untreated. The PV antigen was fused to a human IgG1 hinge region

CC and the resulting fusion protein is useful to treat or diagnose  
 CC pemphigus vulgaris.  
 SQ Sequence 614 AA;

Query Match 24.1%; Score 7; DB 19; Length 614;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 fvdvknk 105  
 |||||  
 QY 22 FVVDKNT 28

RESULT 6  
 ID R30742 standard; Protein; 999 AA.  
 AC R30742;  
 DT 14-JUN-1993 (first entry)  
 DE Human pemphigus vulgaris 130kd antigen.  
 KW Pemphigus vulgaris; skin disease; autoantibodies;  
 KW keratinocyte cell surface antigen; glycoprotein; cell adhesion.  
 OS Homo sapiens.  
 PN US7798918-A.  
 PD 15-DEC-1992.  
 PF 27-NOV-1991; 798918.  
 PR 27-NOV-1991; US-798918.  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.  
 PI Amagai M, Klaus-kovtun V, Stanley JR;  
 DR WPI: 93-067436/08.  
 DR N-PSDB: Q35992.  
 PT DNA encoding pemphigus vulgaris antigen - useful in proteins for diagnostic and therapeutic uses  
 PS Disclosure; Fig 7; 50pp; English.  
 CC This sequence is the pemphigus vulgaris 130kd antigen. The protein and its encoding DNA may be used in the diagnosis and treatment of pemphigus vulgaris. It is thought that the antigen may be a cell adhesion molecule.  
 SQ Sequence 999 AA;

Query Match 24.1%; Score 7; DB 6; Length 999;  
 Best Local Similarity 100.0%; Pred. No. 1.70e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 100 fvdvknk 106  
 |||||  
 QY 22 FVVDKNT 28

RESULT 7  
 ID R65115 standard; peptide; 9 AA.  
 AC R65115;  
 DT 09-OCT-1995 (first entry)  
 DE PAP immunogenic peptide 135-143.  
 KW PAP: immunogenic peptide 135-143; cytotoxic C cells; in vitro activation; cancer; AIDS; bacterial infections; malaria; fungal infections; tuberculosis; hepatitis.  
 OS Homo sapiens.  
 PN W09504817-A.  
 PD 16-FEB-1995.  
 PF 01-AUG-1994; U08672.  
 PR 06-AUG-1993; US-103401.  
 PA (CYTE-) CYTEL CORP.  
 PI Cells E, Kubo R, Serra H, Tsai V, Wentworth P;  
 DR WPI: 95-090895/12.  
 PT In vitro activation of cytotoxic T cells for selected killing of target cells - for treating e.g. cancer, AIDS, hepatitis etc.by incubating them with antigen presenting cells loaded with appropriate immunogenic peptide  
 PS Example 3; Page 35; 53pp; English.  
 CC R65109-R65145 are immunogenic peptides, they are used in a new method for the in vitro activation of cytotoxic T cells (CTC). This is achieved by incubating the CTCs with antigen presenting cells loaded with an appropriate immunogenic peptide (e.g. one of the above peptides). By selecting the peptides used the

CC following diseases and infections can be treated; cancer, AIDS,  
 CC hepatitis, other viral and bacterial infections, malaria and  
 CC tuberculosis.  
 SQ Sequence 9 AA;

Query Match 20.7%; Score 6; DB 13; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.27e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 illwdp 6  
 Qy 9 ILLWDP 14

RESULT 8

ID R65114 standard; peptide; 9 AA.  
 AC R65114;  
 DT 09-OCT-1995 (first entry)  
 DE Hepatitis B virus (HBV) immunogenic peptide 18-27 analog.  
 KW HBV; immunogenic peptide 18-27 analog; cytotoxic C cells;  
 KW in vitro activation; cancer; AIDS; bacterial infections; malaria;  
 KW fungal infections; tuberculosis; hepatitis.  
 OS Homo sapiens.  
 PN WO9504817-A.  
 PD 16-FEB-1995.  
 PF 01-AUG-1994; U08672.  
 PR 06-AUG-1993; US-103401.  
 PA (CYTE-) CYTEL CORP.  
 PI Celis E, Kubo R, Serra H, Tsai V, Wentworth P;  
 DR WPI; 95-090895/12.  
 PT In vitro activation of cytotoxic T cells for selected killing of  
 PT target cells - for treating e.g. cancer, AIDS, hepatitis etc.by  
 PT incubating them with antigen presenting cells loaded with  
 PT appropriate immunogenic peptide  
 PS Example 3; Page 35; 53pp; English.  
 CC R65109-R65145 are immunogenic peptides, they are used in a new  
 CC method for the in vitro activation of cytotoxic T cells (CTC).  
 CC This is achieved by incubating the CTCs with antigen presenting  
 CC cells loaded with an appropriate immunogenic peptide (e.g. one  
 CC of the above peptides). By selecting the peptides used the  
 CC following diseases and infections can be treated; cancer, AIDS,  
 CC hepatitis, other viral and bacterial infections, malaria and  
 CC tuberculosis.  
 SQ Sequence 9 AA;

Query Match 20.7%; Score 6; DB 13; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.27e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 illwdp 6  
 Qy 9 ILLWDP 14

RESULT 9

ID W82588 standard; Protein; 295 AA.  
 AC W82588;  
 DT 01-MAR-1999 (first entry)  
 DE Human ATG-1622 protein.  
 KW ATG-1622; SDF5; human; secreted ligand; 7-transmembrane receptor;  
 KW diagnosis; disease; screening; vaccine; inoculate; treatment; obesity;  
 KW heart disease; hypertension; kidney diseases; insulin resistance;  
 KW lipodystrophy; diabetes; central nervous system; CNS; gene mapping;  
 KW linkage analysis.  
 OS Homo sapiens.  
 PN EP-879887-A1.  
 PD 25-NOV-1998.  
 PF 14-MAY-1998; 303809.  
 PR 13-JUN-1997; US-874156.  
 PR 21-MAY-1997; US-047251.  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PI Hu E, Zhu Y;  
 LR WPI; 98-596879/51.

DR N-PSDB; V69384.  
 PT New human secreted protein ATG-1622 polypeptide and polynucleotide -  
 PT useful as diagnostic reagents and for prevention and treatment of  
 PT Central Nervous System diseases and diabetes  
 PS Claim 11; Page 22-23; 28pp; English.  
 CC This sequence represents the human ATG-1622 protein which is related to  
 CC human secreted ligands for 7-transmembrane receptors and similar to  
 CC murine SDF5. ATG-1622 polypeptides and polynucleotides are useful for  
 CC diagnosing susceptibility to diseases by detecting mutations in the  
 CC ATG-1622 gene and can diagnose diseases associated with ATG-1622  
 CC imbalance. The polypeptides can be used to screen for agonists and  
 CC antagonists which can be used in treatment to activate or inhibit  
 CC ATG-1622 activity. The ATG-1622 polypeptide can be administered directly  
 CC or as a vaccine to inoculate against disease. Diseases which can be  
 CC diagnosed, prevented or treated by the ATG-1622 polypeptide or  
 CC polynucleotides include heart disease, hypertension, kidney diseases,  
 CC obesity, insulin resistance, lipodystrophy, diabetes and central nervous  
 CC system (CNS) diseases. The ATG-1622 polypeptide is also useful for  
 CC mapping the gene to a chromosome, allowing gene inheritance to be  
 CC studied through linkage analysis.  
 SQ Sequence 295 AA;

Query Match 20.7%; Score 6; DB 38; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 2.27e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266  
 Qy 16 GQKQGG 21

RESULT 10

ID W75102 standard; Protein; 295 AA.  
 AC W75102;  
 DT 28-JAN-1999 (first entry)  
 DE Human secreted protein encoded by gene 46 clone HLHDY31.  
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 OS Homo sapiens.  
 FH Key  
 FT Misc-difference 38  
 FT Location/Qualifiers  
 PN WO9839446-A2.  
 PD 11-SEP-1998.  
 PF 06-MAR-1998; U04492.  
 PR 07-MAR-1997; US-038621.  
 PR 07-MAR-1997; US-040161.  
 PR 07-MAR-1997; US-040162.  
 PR 07-MAR-1997; US-040163.  
 PR 07-MAR-1997; US-040333.  
 PR 07-MAR-1997; US-040334.  
 PR 07-MAR-1997; US-040336.  
 PR 07-MAR-1997; US-040626.  
 PR 11-APR-1997; US-043311.  
 PR 11-APR-1997; US-043312.  
 PR 11-APR-1997; US-043313.  
 PR 11-APR-1997; US-043314.  
 PR 11-APR-1997; US-043315.  
 PR 11-APR-1997; US-043568.  
 PR 11-APR-1997; US-043569.  
 PR 11-APR-1997; US-043576.  
 PR 11-APR-1997; US-043578.  
 PR 11-APR-1997; US-043580.  
 PR 11-APR-1997; US-043669.  
 PR 11-APR-1997; US-043670.  
 PR 11-APR-1997; US-043671.  
 PR 11-APR-1997; US-043672.

PR 11-APR-1997; US-043674.  
PR 23-MAY-1997; US-047492.  
PR 23-MAY-1997; US-047500.  
PR 23-MAY-1997; US-047501.  
PR 23-MAY-1997; US-047502.  
PR 23-MAY-1997; US-047503.  
PR 23-MAY-1997; US-047581.  
PR 23-MAY-1997; US-047582.  
PR 23-MAY-1997; US-047583.  
PR 23-MAY-1997; US-047584.  
PR 23-MAY-1997; US-047585.  
PR 23-MAY-1997; US-047586.  
PR 23-MAY-1997; US-047587.  
PR 23-MAY-1997; US-047588.  
PR 23-MAY-1997; US-047589.  
PR 23-MAY-1997; US-047590.  
PR 23-MAY-1997; US-047592.  
PR 23-MAY-1997; US-047593.  
PR 23-MAY-1997; US-047594.  
PR 23-MAY-1997; US-047595.  
PR 23-MAY-1997; US-047596.  
PR 23-MAY-1997; US-047597.  
PR 23-MAY-1997; US-047598.  
PR 23-MAY-1997; US-047599.  
PR 23-MAY-1997; US-047600.  
PR 23-MAY-1997; US-047601.  
PR 23-MAY-1997; US-047612.  
PR 23-MAY-1997; US-047613.  
PR 23-MAY-1997; US-047614.  
PR 23-MAY-1997; US-047615.  
PR 23-MAY-1997; US-047617.  
PR 23-MAY-1997; US-047618.  
PR 23-MAY-1997; US-047632.  
PR 23-MAY-1997; US-047633.  
PR 06-JUN-1997; US-048964.  
PR 06-JUN-1997; US-048974.  
PR 22-AUG-1997; US-056630.  
PR 22-AUG-1997; US-056631.  
PR 22-AUG-1997; US-056632.  
PR 22-AUG-1997; US-056636.  
PR 22-AUG-1997; US-056637.  
PR 22-AUG-1997; US-056652.  
PR 22-AUG-1997; US-056654.  
PR 22-AUG-1997; US-056845.  
PR 22-AUG-1997; US-056846.  
PR 22-AUG-1997; US-056862.  
PR 22-AUG-1997; US-056864.  
PR 22-AUG-1997; US-056872.  
PR 22-AUG-1997; US-056874.  
PR 22-AUG-1997; US-056875.  
PR 22-AUG-1997; US-056876.  
PR 22-AUG-1997; US-056877.  
PR 22-AUG-1997; US-056878.  
PR 22-AUG-1997; US-056879.  
PR 22-AUG-1997; US-056880.  
PR 22-AUG-1997; US-056881.  
PR 22-AUG-1997; US-056882.  
PR 22-AUG-1997; US-056884.  
PR 22-AUG-1997; US-056886.  
PR 22-AUG-1997; US-056887.  
PR 22-AUG-1997; US-056888.  
PR 22-AUG-1997; US-056889.  
PR 22-AUG-1997; US-056892.  
PR 22-AUG-1997; US-056893.  
PR 22-AUG-1997; US-056894.  
PR 22-AUG-1997; US-056903.  
PR 22-AUG-1997; US-056908.  
PR 22-AUG-1997; US-056909.  
PR 22-AUG-1997; US-056910.  
PR 22-AUG-1997; US-056911.  
PR 05-SEP-1997; US-057650.  
PR 05-SEP-1997; US-057761.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,

Feng P, Ferrie AM, Fischer CL, Graves KA, Greene JM, Hu JS,  
Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen Ch,  
Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
WPI: 98-609887/51.  
N-PSDB: V34199.  
New isolated human genes and the secreted polypeptides they encode  
- useful for diagnosis and treatment of e.g. cancers, neurological  
disorders, immune diseases, inflammation of blood disorders  
Claim 1; Page 306-307; 447pp; English.  
This sequence represents a secreted human protein encoded by the gene  
clone detailed in the descriptor line.  
The gene can be used to generate fusion proteins by linking to the gene  
to a human immunoglobulin Fc portion (e.g. V34145) for increasing the  
stability of the fused protein as compared to the human protein only.  
The invention relates to 70 novel genes and their fragments (nucleic acid  
sequences: V34154-V34276; amino acid sequences W5057-W5179) which  
are useful for preventing, treating or ameliorating medical conditions  
e.g. by protein or gene therapy. Also, pathological conditions can be  
diagnosed by determining the amount of the new polypeptides in a sample  
or by determining the presence of mutations in the new polynucleotides.  
Specific uses are described for each of the 70 polynucleotides, based on  
CC which tissues they are most highly expressed in (see V34154 for described  
CC uses).

SQ Sequence 295 AA;  
Query Match 20.7%; Score 6; DB 37; Length 295;  
Best Local Similarity 100.0%; Pred. No. 2.27e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266  
|||||  
Qy 16 GQKQGG 21

RESULT 11  
ID W37814 standard; Protein; 295 AA.  
AC W37814, 1998 (first entry)  
DT 28-AUG-1998  
DE Murine secreted apoptosis-related protein msARPL.  
KW Secreted apoptosis-related protein; SARP; msARPL; mouse;  
KW prostate cancer; breast cancer; diagnosis; gene therapy.  
OS Mus musculus.  
PN W09813493-A2.  
PD 02-APR-1998.  
PF 24-SEP-1997; U17154.  
PR 11-OCT-1996; US-028363.  
PR 24-SEP-1996; US-026603.  
PA (LXRB-) LXR BIOTECHNOLOGY INC.  
PI Melkonian H, Umansky S;  
DR WPI: 98-230704/20.  
N-PSDB: V19112.  
DR New secreted apoptosis-related proteins - useful for modulating  
PT apoptosis, particularly for treatment of prostatic or breast cancer,  
PT also for diagnosis and monitoring of disease  
PS Claim 1; Page 50-51; 101pp; English.  
CC This polypeptide comprises murine secreted apoptosis-related  
CC protein msARPL that modulates apoptosis through cell-cell or  
CC cell-extracellular matrix signalling. Its amino acid sequence was  
CC deduced from a clone (see V19112) obtained from a murine 10T1/2  
CC fibroblast cDNA library. Human SARP1, 2 and 3 (see W37815-17)  
CC proteins are also claimed. msARPL is expressed mainly in heart and  
CC lung with some expression in kidney. SARP polypeptides can be  
CC obtained from recombinant host cells. Antibodies specific for SARP  
CC polypeptides can be used in immunoassays for detecting levels of  
CC expression of SARP, particularly for diagnosis or monitoring of  
CC diseases associated with SARP expression. SARP polypeptides and  
CC nucleic acids can also be used to treat cancers, or more generally  
CC apoptosis-related disease (e.g. infection with HIV or reperfusion  
CC injury), also (not claimed) to prevent apoptosis in cultured cells,  
CC to improve preservation of organs for transplantation, for in situ  
CC preservation for by-pass operations and to treat dermatological  
CC disorders. SARP polypeptides can also be used to identify agents,  
CC potentially useful therapeutically, that modulate the effects of

```
CC SARP on Wnt-frizzled protein interaction.
SQ Sequence 295 AA;

Query Match 20.7%; Score 6; DB 32; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.27e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266
QY 16 GQKQGG 21

RESULT 12
ID W49082 standard; Protein; 295 AA.
AC W49082; 09-NOV-1998 (first entry)
DE Homo sapiens SDF-5 protein.
KW SDF-5; frizzled; osteoarthritis; rheumatoid arthritis;
KW cell formation; proliferation; differentiation; diabetes;
KW pancreatic cancer; wound healing; gene therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..20
FT /note= "signal peptide"
FT WO9835043-A1.
PN 13-AUG-1998.
PD 15-OCT-1997; U18369.
PF 08-MAY-1997; US-848439.
PR 06-FEB-1997; US-796153.
PA (GEM ) GENETICS INST INC.
PI Lavallie ER, Racie LA;
DR WPI: 98-447240/38.
DR N-PSDB: V32930.
PT Isolated DNA encoding human SDF-5 protein - useful for controlling
PT growth, differentiation etc. of cells, particularly of chondrocytes
PT for treatment of arthritis etc., also pancreatic cells
PS Claim 18: Page 57-58; 69pp; English.
CC The sequence is that of human SDF-5, a member of the Frazzled
CC protein family. Cells transfected with a vector containing the
CC sequence are used to regulate genes, particularly pancreatic genes,
CC or in combination with bone morphogenic protein 2 (BMP2), to
CC increase differentiation of progenitor cells into chondrocytes.
CC The protein may be used to treat osteoarthritis, rheumatoid
CC arthritis, or articular cartilage defects, also to increase/inhibit cell
CC formation, growth, differentiation, proliferation and/or maintenance in
CC many other organs or tissues, e.g. for prevention or treatment of
CC pancreatic cancer, diabetes (by inducing de novo formation of islet
CC cells), other tissue defects, also to improve healing of wounds and to
CC increase survival of nervous system cells, e.g. in cases of transplants
CC The coding sequence can be used in gene therapy, and its fragments to
CC detect related mRNA, while the protein is also used to generate
CC antibodies, useful for affinity purification and as immunoassay
CC reagents. Many other potential uses/activities for the gene and its
CC encoded are contemplated but not exemplified, e.g. as cytokines,
CC immuno-suppressants or immunostimulants, regulators of haematopoiesis,
CC as fertility-control agents, haemostatic or thrombolytic agents,
CC anti-inflammatory agents, antimicrobials, modulators of biorhythms
CC and many more.
SQ Sequence 295 AA;

Query Match 20.7%; Score 6; DB 34; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.27e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266
QY 16 GQKQGG 21

RESULT 13
ID W37943 standard; Protein; 295 AA.
AC W37943;
DR 21-AUG-1998 (first entry)
```

```
DE Amino acid sequence of mouse SDF5.
KW Mouse SDF5 protein; bone metabolism abnormality; anxiolytic;
KW immunostimulant.
OS Mus sp.
PN J10136983-A.
PD 26-MAY-1998.
PF 13-NOV-1996; 301666.
PR 13-NOV-1996; JP-301666.
PA (ONOV ) ONO PHARM CO LTD.
DR WPI: 98-355045/31.
DR N-PSDB: V29253; V29254.
PT New polypeptide - used to treat and prevent bone metabolism
PT abnormality and as anxiolytic and immunostimulant
PS Claim 1: Pages 10-11; 17pp; Japanese.
CC This is the amino acid sequence of the mouse SDF5 protein used in
CC the method of the invention, and is encoded by sequences V29253,
CC and V29254. In the method of the invention the polypeptide can be
CC used for the prevention and the treatment of bone metabolism
CC abnormality and as an anxiolytic and immunostimulant.
SQ Sequence 295 AA;

Query Match 20.7%; Score 6; DB 32; Length 295;
Best Local Similarity 100.0%; Pred. No. 2.27e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266
QY 16 GQKQGG 21

RESULT 14
ID W75160 standard; Protein; 296 AA.
AC W75160;
DR 28-JAN-1999 (first entry)
DE Human secreted protein encoded by gene 46 clone HLHDY31.
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 38
FT /label= unknown
FT Misc-difference 296
FT /label= unknown
FT WO9839446-A2.
PD 11-SEP-1998.
PF 06-MAR-1998; U04492.
PR 07-MAR-1997; US-038621.
PR 07-MAR-1997; US-040161.
PR 07-MAR-1997; US-040162.
PR 07-MAR-1997; US-040163.
PR 07-MAR-1997; US-040333.
PR 07-MAR-1997; US-040334.
PR 07-MAR-1997; US-040336.
PR 07-MAR-1997; US-040626.
PR 11-APR-1997; US-043311.
PR 11-APR-1997; US-043312.
PR 11-APR-1997; US-043313.
PR 11-APR-1997; US-043314.
PR 11-APR-1997; US-043315.
PR 11-APR-1997; US-043568.
PR 11-APR-1997; US-043569.
PR 11-APR-1997; US-043576.
PR 11-APR-1997; US-043578.
PR 11-APR-1997; US-043580.
PR 11-APR-1997; US-043669.
PR 11-APR-1997; US-043670.
PR 11-APR-1997; US-043671.
```

```

PR 11-APR-1997; US-043672.
PR 11-APR-1997; US-043674.
PR 23-MAY-1997; US-047492.
PR 23-MAY-1997; US-047500.
PR 23-MAY-1997; US-047501.
PR 23-MAY-1997; US-047502.
PR 23-MAY-1997; US-047503.
PR 23-MAY-1997; US-047581.
PR 23-MAY-1997; US-047582.
PR 23-MAY-1997; US-047583.
PR 23-MAY-1997; US-047584.
PR 23-MAY-1997; US-047585.
PR 23-MAY-1997; US-047586.
PR 23-MAY-1997; US-047587.
PR 23-MAY-1997; US-047588.
PR 23-MAY-1997; US-047589.
PR 23-MAY-1997; US-047590.
PR 23-MAY-1997; US-047592.
PR 23-MAY-1997; US-047593.
PR 23-MAY-1997; US-047594.
PR 23-MAY-1997; US-047595.
PR 23-MAY-1997; US-047596.
PR 23-MAY-1997; US-047597.
PR 23-MAY-1997; US-047598.
PR 23-MAY-1997; US-047599.
PR 23-MAY-1997; US-047600.
PR 23-MAY-1997; US-047601.
PR 23-MAY-1997; US-047612.
PR 23-MAY-1997; US-047613.
PR 23-MAY-1997; US-047614.
PR 23-MAY-1997; US-047615.
PR 23-MAY-1997; US-047617.
PR 23-MAY-1997; US-047618.
PR 23-MAY-1997; US-047632.
PR 23-MAY-1997; US-047633.
PR 06-JUN-1997; US-048964.
PR 06-JUN-1997; US-048974.
PR 22-AUG-1997; US-056630.
PR 22-AUG-1997; US-056631.
PR 22-AUG-1997; US-056632.
PR 22-AUG-1997; US-056636.
PR 22-AUG-1997; US-056637.
PR 22-AUG-1997; US-056682.
PR 22-AUG-1997; US-056684.
PR 22-AUG-1997; US-056845.
PR 22-AUG-1997; US-056862.
PR 22-AUG-1997; US-056864.
PR 22-AUG-1997; US-056872.
PR 22-AUG-1997; US-056874.
PR 22-AUG-1997; US-056875.
PR 22-AUG-1997; US-056876.
PR 22-AUG-1997; US-056877.
PR 22-AUG-1997; US-056878.
PR 22-AUG-1997; US-056879.
PR 22-AUG-1997; US-056880.
PR 22-AUG-1997; US-056881.
PR 22-AUG-1997; US-056882.
PR 22-AUG-1997; US-056884.
PR 22-AUG-1997; US-056886.
PR 22-AUG-1997; US-056887.
PR 22-AUG-1997; US-056888.
PR 22-AUG-1997; US-056889.
PR 22-AUG-1997; US-056892.
PR 22-AUG-1997; US-056893.
PR 22-AUG-1997; US-056894.
PR 22-AUG-1997; US-056903.
PR 22-AUG-1997; US-056908.
PR 22-AUG-1997; US-056909.
PR 22-AUG-1997; US-056910.
PR 22-AUG-1997; US-056911.
PR 05-SEP-1997; US-057650.
PR 05-SEP-1997; US-057761.
PA (HUMA-) HUMAN GENOME SCI INC.

PI Bednarik DP, Brever LA, Carter KC, Duan R, Ebner R, Endress GA,
PI Feng P, Ferrie AM, Fischer CL, Graves KA, Greene JM, Hu JS,
PI Kyaw H, Larleu DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
DR WPI: 98-609887/51.
DR N-PSDB: V34257.
PT New isolated human genes and the secreted polypeptides they encode
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
PS Claim 1; Page 341-342; 447pp; English.
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line.
CC The gene can be used to generate fusion proteins by linking to the gene
CC to a human immunoglobulin Fc portion (e.g. V34145) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 70 novel genes and their fragments (nucleic acid
CC sequences: V34154-V34276; amino acid sequences W75057-W75179) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 70 polynucleotides, based on
CC which tissues they are most highly expressed in (see V34154 for described
CC uses).
SQ Sequence 296 AA;

Query Match 20.7%; Score 6; DB 37; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.27e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 gqkqgg 266
   |||||
QY 16 GQRQGG 21

RESULT 15
ID W75770 standard; Protein; 567 AA.
AC W75770;
DE 21-DEC-1998 (first entry)
DE Human oxidoreductase YTF03.
KW Disintegrin metalloproteinase; proteinase; BS10.55; human.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..22 /label= Sig_peptide
FT Protein 23..567 /label= Mat_protein
FT Region 22 /note= "a mitochondrial localisation motif is near
FT Gln22"
FT Domain 62..78 /note= "putative transmembrane domain"
FT Domain 272..289 /note= "putative transmembrane domain"
FT Peptide 21..44 /note= "antigenic peptide"
FT Peptide 136..157 /note= "antigenic peptide"
FT Peptide 171..189 /note= "antigenic peptide"
FT Peptide 358..390 /note= "antigenic peptide"
FT Peptide 503..567 /note= "antigenic peptide"
PN W09839421-A2.
PD 11-SEP-1998.
PF 06-MAR-1998; U03937.
PR 07-MAR-1997; US-813150.
PA (SCHE ) SCHERING CORP.
PI Dowling LM, Gorman DM, Huffine CF, LeBecque SJE,
PI Liu Y, Mueller CG;
PI WPI: 98-506353/43.
DR N-PSDB: V52605.

```

PT New isolated genes from human dendritic cells - which encode  
PT products having disintegrin-metalloproteinase monoamine oxidase or  
PT GTPase activities.  
PS Claim 2; Page 91-94; 106pp; English.  
CC This is the amino acid sequence of human YFRO3 protein, deduced  
CC from a dendritic cell cDNA sequence (see V52605). The YFRO3 cDNA  
CC is distantly related to a family of enzymes which include monoamine  
CC oxidases, but the translation product does not have the key  
CC residues which covalently bind the FAD cofactor, and so may not  
CC exhibit monoamine oxidase activity. The cellular types which  
CC express messages encoding YFRO3 suggest that signals important in  
CC cell differentiation and development are mediated by them. The  
CC invention provides a process for recombinant production of YFRO3,  
CC host cells, expression vectors and specific antibodies. The  
CC products can be used for detection, diagnosis, development of  
CC therapeutic methods and drug screening.  
SQ Sequence 567 AA;

Query Match 20.7%; Score 6; DB 36; Length 567;  
Best Local Similarity 100.0%; Pred. No. 2.27e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 451 qggfvv 456  
| | | | |  
Qy 19 QGGFVV 24

Search completed: Sat Aug 28 14:55:38 1999  
Job time : 76 secs.



\*\*\*\*\*  
W O E H  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:53:20 1999; MasPar time 5.09 Seconds  
228.474 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-45  
Description: (1-23) from US09049696.pep  
Perfect Score: 29  
Sequence: 1 TWTTPPQILLWDPGKQGGFYVDKNTK 29

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 2.645; Variance 0.403; scale 6.561

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES								
%								
Result	No.	Score	Query	Match	Length	ID	Description	Pred. No.
1	7	24.1	999	1	IJHUG3	desmoglein 3 precursor	5.71e-02	
2	6	20.7	124	2	JC4849	H+-transporting ATPase	3.54e+00	
3	6	20.7	124	2	A53055	H+-transporting ATPase	3.54e+00	
4	6	20.7	154	2	S48417	hypothetical protein	3.54e+00	
5	6	20.7	206	2	S68597	NADPH dehydrogenase (	3.54e+00	
6	6	20.7	264	2	A71708	acyl-[acyl-carrier-pr	3.54e+00	
7	6	20.7	295	2	JE0174	frizzled protein-2 -	3.54e+00	
8	6	20.7	325	2	F71066	hypothetical protein	3.54e+00	
9	6	20.7	328	2	G70357	flagellar switch prot	3.54e+00	
10	6	20.7	555	2	C70177	beta-glucosidase homo	3.54e+00	
11	6	20.7	1116	2	B70476	hypothetical protein	3.54e+00	
12	6	20.7	1323	2	S27224	N-methyl-D-aspartate	3.54e+00	
13	6	20.7	1323	2	F78557	N-methyl-D-aspartate	3.54e+00	
14	6	20.7	1356	2	C45219	N-methyl-D-aspartate	3.54e+00	
15	6	20.7	2326	2	B47447	calcium channel prote	3.54e+00	
16	6	20.7	3169	2	T00296	toxin B - Escherichia	3.54e+00	
17	5	17.2	15	2	A56891	gamma 1 gliadin - whe	1.44e+02	
18	5	17.2	15	2	B56891	gamma 2 gliadin - whe	1.44e+02	
19	5	17.2	18	2	S51255	gamma2-gliadin p25-27	1.44e+02	
20	5	17.2	22	2	A34928	myosin heavy chain, b	1.44e+02	
21	5	17.2	30	2	S52126	gamma3-gliadin p25-27	1.44e+02	
22	5	17.2	66	1	NT5R2N	neurotoxin 2 - scorpi	1.44e+02	
23	5	17.2	68	2	F70214	lipoprotein homolog -	1.44e+02	

myosin heavy chain, n 1.44e+02  
GTP-binding regulator 1.44e+02  
hypothetical protein 1.44e+02  
conserved hypothetical 1.44e+02  
transposase IS240-A h 1.44e+02  
hypothetical protein 1.44e+02  
flagellar protein fl 1.44e+02  
cytochrome c3 precurs 1.44e+02  
T-cell receptor alpha 1.44e+02  
16K lactose-binding 1 1.44e+02  
yhcV homolog - Methan 1.44e+02  
hypothetical protein 1.44e+02  
icmX protein precurs 1.44e+02  
low-temperature regul 1.44e+02  
low-temperature regul 1.44e+02  
TSC-22 protein - mous 1.44e+02  
transposase IS240-A h 1.44e+02  
TGF beta-stimulated c 1.44e+02  
hypothetical protein 1.44e+02  
hypothetical protein 1.44e+02  
STH-2 protein - potat 1.44e+02  
hypothetical protein 1.44e+02  
hypothetical protein 1.44e+02  
myosin light chain-2 1.44e+02  
lipoprotein signal pe 1.44e+02  
ORF2 protein - Chlore 1.44e+02  
probable moab2 protei 1.44e+02  
zein Zcl - maize 1.44e+02  
conserved hypothetical 1.44e+02  
hypothetical protein 1.44e+02  
repressor for cytochr 1.44e+02  
genome polypotein - 1.44e+02  
hypothetical protein 1.44e+02  
nopaline - Agrobacter 1.44e+02  
hypothetical protein 1.44e+02  
NADPH dehydrogenase ( 1.44e+02  
hypothetical protein 1.44e+02  
hypothetical protein 1.44e+02  
conserved hypothetical 1.44e+02  
CMRF-35 antigen - hum 1.44e+02  
hypothetical protein 1.44e+02  
smooth muscle myosin 1.44e+02  
probable transcriptio 1.44e+02  
protein-L-isoaspartat 1.44e+02  
phosphate regulon reg 1.44e+02  
site-specific DNA end 1.44e+02  
membrane fusion prote 1.44e+02  
probable efflux trans 1.44e+02  
resiniferatoxin-bindi 1.44e+02  
microfilarial sheath 1.44e+02  
hypothetical protein 1.44e+02  
hypothetical protein 1.44e+02  
nitrate ABC transport 1.44e+02  
ribosomal protein S3, 1.44e+02  
ribosomal protein S3, 1.44e+02  
hypothetical protein 1.44e+02  
ribosomal protein S1 1.44e+02  
ribosomal protein S1 1.44e+02  
deoxycytidylate 5-hyd 1.44e+02  
peptidyl-prolyl cis-t 1.44e+02  
deoxycytidylate 5-hyd 1.44e+02  
hypothetical protein 1.44e+02  
probable NADH dehydro 1.44e+02  
o-antigen export syst 1.44e+02  
H repeat-associated p 1.44e+02  
hypothetical protein 1.44e+02  
ferric enterobactin t 1.44e+02  
MUC1 protein - human 1.44e+02

24 5 17.2 86 2 PN0013  
25 5 17.2 88 2 S41148  
26 5 17.2 89 2 S76803  
27 5 17.2 92 2 B69150  
28 5 17.2 93 2 G69361  
29 5 17.2 114 2 T03081  
30 5 17.2 121 2 C64956  
31 5 17.2 129 1 C6DV3  
32 5 17.2 130 2 A25020  
33 5 17.2 135 2 A42812  
34 5 17.2 138 2 B64415  
35 5 17.2 139 2 S6128  
36 5 17.2 140 2 S61385  
37 5 17.2 142 2 JQ2279  
38 5 17.2 142 2 JQ2280  
39 5 17.2 143 2 S23255  
40 5 17.2 143 2 B69267  
41 5 17.2 144 2 JC4813  
42 5 17.2 151 2 B69183  
43 5 17.2 152 2 I41310  
44 5 17.2 153 2 S63255  
45 5 17.2 155 2 S35161  
46 5 17.2 159 2 A32048  
47 5 17.2 159 2 E71208  
48 5 17.2 163 2 A53718  
49 5 17.2 164 1 ZPECL  
50 5 17.2 171 2 H64107  
51 5 17.2 173 2 B46352  
52 5 17.2 181 2 D70821  
53 5 17.2 183 2 B29017  
54 5 17.2 186 2 F69850  
55 5 17.2 187 2 S74987  
56 5 17.2 192 2 A42116  
57 5 17.2 193 2 S35630  
58 5 17.2 196 2 S42038  
59 5 17.2 204 2 S41893  
60 5 17.2 204 2 E65183  
61 5 17.2 206 2 JC2070  
62 5 17.2 208 2 S46510  
63 5 17.2 216 2 D70554  
64 5 17.2 219 2 F70474  
65 5 17.2 224 2 I37243  
66 5 17.2 225 2 S57810  
67 5 17.2 226 2 I65769  
68 5 17.2 227 2 A64790  
69 5 17.2 229 2 E70471  
70 5 17.2 230 2 S66344  
71 5 17.2 231 2 F64120  
72 5 17.2 234 2 S56763  
73 5 17.2 234 2 F64595  
74 5 17.2 234 2 G1918  
75 5 17.2 235 2 S33448  
76 5 17.2 237 2 S46964  
77 5 17.2 237 2 A64532  
78 5 17.2 237 2 H19174  
79 5 17.2 240 2 S75061  
80 5 17.2 243 2 G69260  
81 5 17.2 243 2 S41170  
82 5 17.2 243 1 R3R13  
83 5 17.2 245 2 F64465  
84 5 17.2 246 1 R3XL3A  
85 5 17.2 246 2 I51635  
86 5 17.2 246 1 S2BPT6  
87 5 17.2 246 1 S2BPT6  
88 5 17.2 246 2 S74559  
89 5 17.2 246 1 S2BPT2  
90 5 17.2 246 2 H39999  
91 5 17.2 249 2 T03189  
92 5 17.2 252 2 E71707  
93 5 17.2 253 2 A64806  
94 5 17.2 253 2 S6719  
95 5 17.2 254 2 A64311  
96 5 17.2 255 2 S48146

97	5	17.2	256	2	B71264	hypothetical protein	1.44e+02	170	5	17.2	392	2	S55971	hypothetical protein	1.44e+02
98	5	17.2	256	2	JC4627	fibroblast growth fac	1.44e+02	171	5	17.2	392	2	S74790	hypothetical protein	1.44e+02
99	5	17.2	257	2	E64480	cytochrome-c3 hydroge	1.44e+02	172	5	17.2	393	2	S59499	cellulase egli - smut	1.44e+02
100	5	17.2	265	2	T02138	hypothetical protein	1.44e+02	173	5	17.2	393	1	VHN2PM	nucleocapsid protein	1.44e+02
101	5	17.2	267	2	S15486	probable ABC-type tra	1.44e+02	174	5	17.2	395	2	S74355	hypothetical protein	1.44e+02
102	5	17.2	268	2	S31010	gene 65 protein - Myc	1.44e+02	175	5	17.2	402	2	S11772	polygalacturonase (EC	1.44e+02
103	5	17.2	269	2	E69381	hypothetical protein	1.44e+02	176	5	17.2	406	2	JC1219	polygalacturonase (EC	1.44e+02
104	5	17.2	270	2	I46620	Sus scrofa ppointle	1.44e+02	177	5	17.2	408	2	A70594	probable manA protein	1.44e+02
105	5	17.2	270	1	S10532	interleukin-1 alpha p	1.44e+02	178	5	17.2	412	2	G30315	polyferredoxin - Meth	1.44e+02
106	5	17.2	273	1	J02390	myb-related protein 2	1.44e+02	179	5	17.2	412	2	H69017	polyferredoxin (MvHB)	1.44e+02
107	5	17.2	273	2	S81332	gvpL protein - Haloba	1.44e+02	180	5	17.2	414	2	A38331	LEP100 protein precu	1.44e+02
108	5	17.2	276	2	S74640	adenylsulfate kinas	1.44e+02	181	5	17.2	416	2	E71474	probable muramoylalan	1.44e+02
109	5	17.2	280	2	A33189	apoptosis suppressor	1.44e+02	182	5	17.2	420	2	A64220	hypothetical protein	1.44e+02
110	5	17.2	280	2	H69314	conserved hypothetica	1.44e+02	183	5	17.2	423	2	S41734	glycine amidinotransf	1.44e+02
111	5	17.2	285	2	H64805	H repeat-associated p	1.44e+02	184	5	17.2	429	2	S27793	hypothetical protein	1.44e+02
112	5	17.2	286	2	E64931	hypothetical protein	1.44e+02	185	5	17.2	429	2	JC4986	site-specific DNA-met	1.44e+02
113	5	17.2	289	2	S75197	ABC-type transport pr	1.44e+02	186	5	17.2	430	2	JC2301	hypothetical 47.8K pr	1.44e+02
114	5	17.2	289	1	SYECSA	succinate-CoA ligase	1.44e+02	187	5	17.2	436	2	A42440	estrogen-regulated pr	1.44e+02
115	5	17.2	290	2	T02300	GTP-binding regulator	1.44e+02	188	5	17.2	438	2	G71175	hypothetical protein	1.44e+02
116	5	17.2	291	2	I38167	gene C6.1A protein -	1.44e+02	189	5	17.2	441	2	S05956	tubulin beta-2 chain	1.44e+02
117	5	17.2	291	1	EW7G	gamma-gliadin B precu	1.44e+02	190	5	17.2	444	2	I40417	glycerol-3-phosphate	1.44e+02
118	5	17.2	293	2	D64189	succinate-CoA ligase	1.44e+02	191	5	17.2	445	2	S73956	Mg148 homolog VXPSP7	1.44e+02
119	5	17.2	298	2	G71841	hypothetical protein	1.44e+02	192	5	17.2	447	2	I38975	nuclear orphan recept	1.44e+02
120	5	17.2	298	2	B64674	conserved hypothetica	1.44e+02	193	5	17.2	447	2	B64894	hypothetical protein	1.44e+02
121	5	17.2	301	2	S31935	ADP,ATP carrier prote	1.44e+02	194	5	17.2	448	2	A38445	EV12B protein precurs	1.44e+02
122	5	17.2	302	2	JA0153	gamma-gliadin precurs	1.44e+02	195	5	17.2	450	2	B36810	hypothetical protein	1.44e+02
123	5	17.2	304	1	S71285	myb-related protein,	1.44e+02	196	5	17.2	451	2	B70058	conserved hypothetica	1.44e+02
124	5	17.2	307	1	G8FF3	salivary glue protein	1.44e+02	197	5	17.2	452	2	G71416	probable glucosyltran	1.44e+02
125	5	17.2	308	2	S57377	probable membrane pro	1.44e+02	198	5	17.2	460	2	G69917	hypothetical protein	1.44e+02
126	5	17.2	309	2	B64041	hypothetical protein	1.44e+02	199	5	17.2	460	2	F70975	nitrite extrusion pro	1.44e+02
127	5	17.2	311	2	H71416	hypothetical protein	1.44e+02	200	5	17.2	463	1	GRECNK	hypothetical protein	1.44e+02
128	5	17.2	312	2	S73494	L-lactate dehydrogena	1.44e+02	201	5	17.2	463	2	F70627	probable narU protein	1.44e+02
129	5	17.2	312	2	H64250	L-lactate dehydrogena	1.44e+02	202	5	17.2	467	2	S59430	hypothetical protein	1.44e+02
130	5	17.2	315	2	S76696	prohibitin homolog YG	1.44e+02	203	5	17.2	467	2	S64450	probable membrane pro	1.44e+02
131	5	17.2	317	3	T00146	hypothetical protein	1.44e+02	204	5	17.2	467	2	S61141	probable membrane pro	1.44e+02
132	5	17.2	318	2	B64900	hypothetical protein	1.44e+02	205	5	17.2	469	2	A64411	hypothetical protein	1.44e+02
133	5	17.2	319	2	S44172	phenoxylbenzoate dioxy	1.44e+02	206	5	17.2	469	2	S64163	3-isopropylmalate deh	1.44e+02
134	5	17.2	324	2	S77342	probable pyruvate deh	1.44e+02	207	5	17.2	475	2	S54511	hypothetical protein	1.44e+02
135	5	17.2	326	2	T02340	GTP-binding regulator	1.44e+02	208	5	17.2	476	2	S71360	noeA protein - Rhizob	1.44e+02
136	5	17.2	327	2	J50402	gamma-gliadin precurs	1.44e+02	209	5	17.2	477	2	S01696	gene P3 protein - hum	1.44e+02
137	5	17.2	331	2	A42602	recombination protein	1.44e+02	210	5	17.2	478	2	S33886	DNA-directed RNA poly	1.44e+02
138	5	17.2	332	2	J70585	minor endoglucanase (	1.44e+02	211	5	17.2	478	2	D64230	NADH oxidase (nox) ho	1.44e+02
139	5	17.2	332	2	A45710	BRRF2 gene homolog -	1.44e+02	212	5	17.2	478	2	A71417	hypothetical protein	1.44e+02
140	5	17.2	333	1	SVRTSA	succinate-CoA ligase	1.44e+02	213	5	17.2	479	2	F70680	probable obg - Mycoba	1.44e+02
141	5	17.2	336	2	E69845	thiamin biosynthesis	1.44e+02	214	5	17.2	479	2	S52705	probable membrane pro	1.44e+02
142	5	17.2	338	2	P43655	calcium-binding prote	1.44e+02	215	5	17.2	482	2	S76376	hypothetical protein	1.44e+02
143	5	17.2	338	2	E71872	hypothetical protein	1.44e+02	216	5	17.2	488	1	RKKKLA	ribulose-bisphosphate	1.44e+02
144	5	17.2	342	2	S23764	polygalacturonase-inh	1.44e+02	217	5	17.2	495	2	C71410	hypothetical protein	1.44e+02
145	5	17.2	344	1	YFBSA	phenylalanine--tRNA l	1.44e+02	218	5	17.2	495	2	S16447	NADH dehydrogenase (u	1.44e+02
146	5	17.2	345	2	S30579	succinate-CoA ligase	1.44e+02	219	5	17.2	495	2	S25942	O-succinylbenzoate-C	1.44e+02
147	5	17.2	349	2	S51267	probable galactosyltr	1.44e+02	220	5	17.2	497	2	S74404	cytochrome P450 2K1 -	1.44e+02
148	5	17.2	354	2	C89755	hypothetical protein	1.44e+02	221	5	17.2	501	2	A38650	myosin heavy chain, e	1.44e+02
149	5	17.2	355	2	C69308	immunogenic protein (	1.44e+02	222	5	17.2	501	2	S76563	hypothetical protein	1.44e+02
150	5	17.2	356	2	I39510	carboxylesterase (EC	1.44e+02	223	5	17.2	502	2	B42030	alpha-globin transcri	1.44e+02
151	5	17.2	357	2	S61415	naringenin 3-dioxygen	1.44e+02	224	5	17.2	504	1	VCPV3A	coat protein - adeno	1.44e+02
152	5	17.2	357	2	S71772	naringenin 3-dioxygen	1.44e+02	225	5	17.2	504	2	S45644	cytochrome P450 2K1 -	1.44e+02
153	5	17.2	360	2	D64955	probable 1-aminocyclo	1.44e+02	226	5	17.2	510	2	S55204	hypothetical protein	1.44e+02
154	5	17.2	360	2	G71444	hypothetical protein	1.44e+02	227	5	17.2	510	2	S68116	4-aminobutyrate trans	1.44e+02
155	5	17.2	361	2	C71242	hypothetical protein	1.44e+02	228	5	17.2	512	2	A55206	sucrose/fructan hydro	1.44e+02
156	5	17.2	362	2	H69536	cytochrome oxidase,su	1.44e+02	229	5	17.2	515	2	S10572	epithelial tumor anti	1.44e+02
157	5	17.2	363	2	B64807	abrB protein - Escher	1.44e+02	230	5	17.2	515	1	VHN2P3	nucleocapsid protein	1.44e+02
158	5	17.2	364	2	I45915	interstitial retinol	1.44e+02	231	5	17.2	520	2	JC4699	cadmium, zinc, cobalt	1.44e+02
159	5	17.2	365	1	O8ECSA	sulfate transport ATP	1.44e+02	232	5	17.2	520	2	B33830	cation efflux system	1.44e+02
160	5	17.2	366	2	S13363	actin modulator prote	1.44e+02	233	5	17.2	522	2	S12042	glucose transport pro	1.44e+02
161	5	17.2	369	2	A39157	probable RNA helicase	1.44e+02	234	5	17.2	527	2	G64818	probable membrane pro	1.44e+02
162	5	17.2	372	2	S34803	exoh protein - Rhizob	1.44e+02	235	5	17.2	531	2	S62796	probable lipoprotein	1.44e+02
163	5	17.2	374	2	A46352	ORF1 protein - Chloze	1.44e+02	236	5	17.2	538	1	ROMVIM	gag polyprotein - Mol	1.44e+02
164	5	17.2	378	2	F64300	formate dehydrogenase	1.44e+02	237	5	17.2	539	2	A64441	O-sialoglycoprotein e	1.44e+02
165	5	17.2	378	2	G64898	H repeat-associated p	1.44e+02	238	5	17.2	547	2	S53920	SNR1 protein - yeast	1.44e+02
166	5	17.2	378	2	S47703	H repeat-associated p	1.44e+02	239	5	17.2	558	2	F71175	hypothetical protein	1.44e+02
167	5	17.2	382	1	I39848	cell division initiat	1.44e+02	240	5	17.2	564	2	F71975	methyl-accepting chem	1.44e+02
168	5	17.2	386	2	S33965	hypothetical protein	1.44e+02	241	5	17.2	565	2	G64532	hypothetical protein	1.44e+02
169	5	17.2	391	2	S44161	L-arginine--glycine a	1.44e+02	242	5	17.2	565	2	S73854	hypothetical protein	1.44e+02

243	5	17.2	558	2	G02753	testis specific basic	1.44e+02	316	5	17.2	900	2	S66264	55.11 protein homolog	1.44e+02
244	5	17.2	575	2	T01552	hypothetical protein	1.44e+02	317	5	17.2	907	1	Q0BE21	membrane antigen gp35	1.44e+02
245	5	17.2	575	2	S12525	prtd protein - Erwini	1.44e+02	318	5	17.2	911	2	S70911	transferrin-binding p	1.44e+02
246	5	17.2	581	2	A05204	hypothetical protein	1.44e+02	319	5	17.2	912	2	C64107	transferrin-binding p	1.44e+02
247	5	17.2	583	2	J39428	alcam - human	1.44e+02	320	5	17.2	912	2	S70901	transferrin-binding p	1.44e+02
248	5	17.2	587	2	JH0464	DM-GRASP precursor -	1.44e+02	321	5	17.2	913	3	JG0168	gob-5 protein - Mouse	1.44e+02
249	5	17.2	587	1	B44276	coat protein Vp1 - pa	1.44e+02	322	5	17.2	913	1	VGEBFS	glycoprotein gII prec	1.44e+02
250	5	17.2	587	1	RSKVAO	L-ascorbate oxidase (	1.44e+02	323	5	17.2	914	2	S70906	transferrin-binding p	1.44e+02
251	5	17.2	588	2	JH0506	adhesion molecule scl	1.44e+02	324	5	17.2	928	1	VGEBEG	glycoprotein gI precu	1.44e+02
252	5	17.2	588	2	A45254	surface glycoprotein	1.44e+02	325	5	17.2	932	2	S74597	mutL protein homolog	1.44e+02
253	5	17.2	589	2	S59370	alpha-glucosidase hom	1.44e+02	326	5	17.2	932	1	VGEBEC	glycoprotein gI precu	1.44e+02
254	5	17.2	591	2	G01586	p67 - human	1.44e+02	327	5	17.2	943	2	S59317	DIP2 protein - yeast	1.44e+02
255	5	17.2	592	2	D70836	hypothetical protein	1.44e+02	328	5	17.2	948	2	S73549	exonuclease ABC chai	1.44e+02
256	5	17.2	606	2	B69805	conserved hypothetical	1.44e+02	329	5	17.2	959	2	S32016	flagellum-associated	1.44e+02
257	5	17.2	610	2	C70126	DNA mismatch repair p	1.44e+02	330	5	17.2	979	2	A70848	probable membrane pro	1.44e+02
258	5	17.2	612	1	B39019	glucose dehydrogenase	1.44e+02	331	5	17.2	980	2	A38523	genome polyprotein -	1.44e+02
259	5	17.2	637	2	A45777	cell division control	1.44e+02	332	5	17.2	997	2	S63064	probable membrane pro	1.44e+02
260	5	17.2	637	1	A43350	formate-tetrahydrofo	1.44e+02	333	5	17.2	1002	2	JH0292	FUN12 protein - yeast	1.44e+02
261	5	17.2	639	2	S03547	hypothetical protein	1.44e+02	334	5	17.2	1003	2	JH0823	FL-160-2 protein - Tr	1.44e+02
262	5	17.2	639	1	WWV070	70K protein - potato	1.44e+02	335	5	17.2	1024	1	GBEC	beta-galactosidase (E	1.44e+02
263	5	17.2	643	2	S69681	hypothetical protein	1.44e+02	336	5	17.2	1025	2	C64974	hypothetical protein	1.44e+02
264	5	17.2	656	2	S38787	CHML protein - human	1.44e+02	337	5	17.2	1033	2	S02168	type I site-specific	1.44e+02
265	5	17.2	656	1	ISBTY3	DNA topoisomerase (EC	1.44e+02	338	5	17.2	1034	2	S63536	aminomethyltransferas	1.44e+02
266	5	17.2	660	2	S73597	lactococcin transport	1.44e+02	339	5	17.2	1037	2	S63535	aminomethyltransferas	1.44e+02
267	5	17.2	660	2	B64243	lactococcin transport	1.44e+02	340	5	17.2	1037	2	S40216	p protein - Flavaria	1.44e+02
268	5	17.2	682	2	S42799	garp precursor - huma	1.44e+02	341	5	17.2	1040	2	B64974	hypothetical protein	1.44e+02
269	5	17.2	666	2	S50452	hypothetical protein	1.44e+02	342	5	17.2	1063	2	T00624	probable endo-1,4-bet	1.44e+02
270	5	17.2	669	2	S65551	factor H - bovine (fr	1.44e+02	343	5	17.2	1090	2	S11823	alpha-dextrin endo-1,	1.44e+02
271	5	17.2	673	2	S36747	acetylcholine regulat	1.44e+02	344	5	17.2	1120	2	S46111	probable membrane pro	1.44e+02
272	5	17.2	677	2	A27286	levanase (EC 3.2.1.65	1.44e+02	345	5	17.2	1144	2	A38968	PI-like adhesin precu	1.44e+02
273	5	17.2	687	2	A46336	glycine--tRNA ligase	1.44e+02	346	5	17.2	1147	1	MWAXIB	myosin heavy chain IB	1.44e+02
274	5	17.2	689	2	S17875	polynucleotide adenyl	1.44e+02	347	5	17.2	1163	1	RWHUIC	cell surface glycopro	1.44e+02
275	5	17.2	698	2	S49206	G1 cyclin CLN1 - yeas	1.44e+02	348	5	17.2	1172	2	T00065	hypothetical protein	1.44e+02
276	5	17.2	707	2	A28438	lactoferrin precursor	1.44e+02	349	5	17.2	1184	2	A55184	fibulin-2 precursor -	1.44e+02
277	5	17.2	708	2	S52317	quinohomoprotein etha	1.44e+02	350	5	17.2	1216	2	A55620	apical endosomal prot	1.44e+02
278	5	17.2	709	2	S40926	hypothetical protein	1.44e+02	351	5	17.2	1221	2	A49457	fibulin-2 precursor -	1.44e+02
279	5	17.2	710	2	S68161	oligopeptide transpor	1.44e+02	352	5	17.2	1231	1	NBHUH	complement factor H p	1.44e+02
280	5	17.2	710	1	Q0BE22	membrane antigen gp22	1.44e+02	353	5	17.2	1244	2	S25327	cytoskeleton assembly	1.44e+02
281	5	17.2	710	2	S72497	oligopeptide transpor	1.44e+02	354	5	17.2	1254	2	S27338	p-glycoprotein C - Ca	1.44e+02
282	5	17.2	711	1	TFHUL	lactotransferrin prec	1.44e+02	355	5	17.2	1255	2	B35175	episialin B - human	1.44e+02
283	5	17.2	716	2	G01627	androgen receptor 1 -	1.44e+02	356	5	17.2	1264	2	A35175	episialin A - human	1.44e+02
284	5	17.2	720	2	J01576	ABI3 protein - Arabid	1.44e+02	357	5	17.2	1279	1	DVRH3	multidrug resistance	1.44e+02
285	5	17.2	738	2	S52737	NADH dehydrogenase (u	1.44e+02	358	5	17.2	1286	1	RJBOP	interphotoreceptor re	1.44e+02
286	5	17.2	739	2	S18642	polynucleotide adenyl	1.44e+02	359	5	17.2	1290	2	JC5473	dextranucrase (EC 2.	1.44e+02
287	5	17.2	740	2	S17925	polynucleotide adenyl	1.44e+02	360	5	17.2	1293	2	T01512	hypothetical protein	1.44e+02
288	5	17.2	744	2	I51022	B24 protein - eastern	1.44e+02	361	5	17.2	1295	2	A35886	polymorphic epithelia	1.44e+02
289	5	17.2	745	2	B70017	xanthine dehydrogenas	1.44e+02	362	5	17.2	1335	2	A35887	mucin precursor, panc	1.44e+02
290	5	17.2	753	2	F69338	pyruvate,water dikina	1.44e+02	363	5	17.2	1380	2	S57150	ZMS1 protein - yeast	1.44e+02
291	5	17.2	754	2	T00393	Nijmegen breakage syn	1.44e+02	364	5	17.2	1407	1	BWBYM1	TATA box-binding prot	1.44e+02
292	5	17.2	756	2	S47656	TMDC II protein - cra	1.44e+02	365	5	17.2	1520	2	G69634	glutamate synthase (l	1.44e+02
293	5	17.2	763	2	A29562	PR1 protein - yeast	1.44e+02	366	5	17.2	1624	2	C70867	probable Helix-turn-h	1.44e+02
294	5	17.2	768	2	T00073	hypothetical protein	1.44e+02	367	5	17.2	1662	2	T01893	hypothetical protein	1.44e+02
295	5	17.2	780	2	T00366	hypothetical protein	1.44e+02	368	5	17.2	1742	2	S76110	hypothetical protein	1.44e+02
296	5	17.2	781	2	S51592	XynB precursor - Rumi	1.44e+02	369	5	17.2	1774	2	B56101	collagen alpha 1(XVII	1.44e+02
297	5	17.2	794	2	S73877	probable lipoprotein	1.44e+02	370	5	17.2	1776	1	RWPYIM	genome polyprotein -	1.44e+02
298	5	17.2	806	2	I51589	PI homolog - African	1.44e+02	371	5	17.2	1839	1	OYBYK	adenylate cyclase (EC	1.44e+02
299	5	17.2	807	2	I51685	MCM3 - African clawed	1.44e+02	372	5	17.2	1867	2	S22775	MOT1 protein - yeast	1.44e+02
300	5	17.2	815	2	B30843	glutenin high molecu	1.44e+02	373	5	17.2	1948	2	S00485	gene 11-1 protein pre	1.44e+02
301	5	17.2	815	2	JN0589	glutenin, high-molecu	1.44e+02	374	5	17.2	1955	1	AGCH	agrin precursor - chl	1.44e+02
302	5	17.2	819	1	TNBE11	91.8K alpha trans-ind	1.44e+02	375	5	17.2	1999	1	S21801	myosin heavy chain, n	1.44e+02
303	5	17.2	820	2	A45178	Raf protein kinase ho	1.44e+02	376	5	17.2	2206	2	G71611	hypothetical protein	1.44e+02
304	5	17.2	827	2	E64607	DNA gyrase, sub A - H	1.44e+02	377	5	17.2	2291	2	S11238	polymetase - Berne vi	1.44e+02
305	5	17.2	831	2	E71906	DNA gyrase chain A -	1.44e+02	378	5	17.2	2376	2	S48405	probable membrane pro	1.44e+02
306	5	17.2	838	2	A70043	pyruvate,water dikina	1.44e+02	379	5	17.2	2505	1	XYRTFA	fatty-acid synthase (	1.44e+02
307	5	17.2	845	2	A40016	matrin 3 - rat	1.44e+02	380	5	17.2	2567	2	A49551	filamin, Mueller cell	1.44e+02
308	5	17.2	848	2	E64246	exonuclease ABC chai	1.44e+02	381	5	17.2	2723	2	T03221	probable polyketide s	1.44e+02
309	5	17.2	859	1	S06418	3',5'-cyclic-GMP phos	1.44e+02	382	5	17.2	2844	2	S28291	hypothetical protein	1.44e+02
310	5	17.2	859	2	B34611	3',5'-cyclic-GMP phos	1.44e+02	383	5	17.2	2897	2	B48666	cell proliferation an	1.44e+02
311	5	17.2	859	2	S13030	3',5'-cyclic-GMP phos	1.44e+02	384	5	17.2	3256	2	A48666	cell proliferation an	1.44e+02
312	5	17.2	868	1	VGEB31	glycoprotein B - huma	1.44e+02	385	5	17.2	3433	1	GNWVKV	genome polyprotein -	1.44e+02
313	5	17.2	869	2	PH918	FL-160-3 protein - Tr	1.44e+02	386	5	17.2	4639	2	A54794	dynein heavy chain, c	1.44e+02
314	5	17.2	876	2	T00540	probable protein kina	1.44e+02	387	5	17.2	4644	1	A38905	dynein heavy chain, c	1.44e+02
315	5	17.2	886	2	S29605	glycoprotein 350/220	1.44e+02	388	5	17.2	4725	1	A44357	dynein heavy chain, c	1.44e+02

389	5	17.2	6669	2	S53024	nebulin, skeletal mus	1.44e+02	462	4	13.8	218	2	A42737	sodium channel beta 1	3.31e+03
390	5	17.2	26926	1	I38344	titin, cardiac muscle	1.44e+02	463	4	13.8	218	2	S58769	brain-specific protei	3.31e+03
391	4	13.8	19	2	G61458	Ig lambda chain V-II	3.31e+03	464	4	13.8	218	2	G70438	hypothetical protein	3.31e+03
392	4	13.8	13	2	A37968	neural surface protei	3.31e+03	465	4	13.8	219	2	E70348	UDP-3-O-(3-hydroxymy	3.31e+03
393	31	14.276	31	2	I46276	hemoglobin beta-x cha	3.31e+03	466	4	13.8	220	2	H71360	probable thiamine ABC	3.31e+03
394	4	13.8	33	2	A58533	CD41 homolog - channe	3.31e+03	467	4	13.8	220	2	F71052	probable adenylate ki	3.31e+03
395	4	13.8	34	2	PC4275	dentin sialoprotein -	3.31e+03	468	4	13.8	220	2	S35075	class II histocompati	3.31e+03
396	4	13.8	39	2	A31763	neural cell adhesion	3.31e+03	469	4	13.8	221	2	B70348	hypothetical protein	3.31e+03
397	4	13.8	54	2	S12504	mb-1 protein - mouse	3.31e+03	470	4	13.8	223	2	B70372	flagellar protein Flg	3.31e+03
398	4	13.8	58	2	S78418	ribosomal protein L32	3.31e+03	471	4	13.8	223	2	F70473	hypothetical protein	3.31e+03
399	4	13.8	58	2	A53313	heparin-binding epide	3.31e+03	472	4	13.8	228	2	S20416	Asf-56 protein - pig	3.31e+03
400	4	13.8	66	3	JN0423	esterase S - fruit fl	3.31e+03	473	4	13.8	229	2	JC5409	platelet-activating f	3.31e+03
401	4	13.8	70	2	PC2063	neurite outgrowth mol	3.31e+03	474	4	13.8	231	2	S28186	achaete-scute locus p	3.31e+03
402	4	13.8	75	2	S22461	hypothetical protein	3.31e+03	475	4	13.8	232	2	E70372	hypothetical protein	3.31e+03
403	4	13.8	85	2	S29576	Ig light chain - rain	3.31e+03	476	4	13.8	233	2	S11563	probable MASH-2 prote	3.31e+03
404	4	13.8	86	3	A41772	maxadilan - Lutomyia	3.31e+03	477	4	13.8	234	2	S01320	Ig kappa chain precu	3.31e+03
405	4	13.8	86	2	S06031	hypothetical protein	3.31e+03	478	4	13.8	234	2	JC4050	Max-associated protei	3.31e+03
406	4	13.8	92	2	S06033	hypothetical protein	3.31e+03	479	4	13.8	235	2	G71259	probable ABC transpor	3.31e+03
407	4	13.8	96	2	S43974	p68 protein - human	3.31e+03	480	4	13.8	235	2	S20000	Ig light chain precu	3.31e+03
408	4	13.8	99	2	B53116	Ig epsilon chain C re	3.31e+03	481	4	13.8	237	2	H71250	probable ATP-depend	3.31e+03
409	4	13.8	102	2	S52241	xiGL2 protein - Afri	3.31e+03	482	4	13.8	237	2	B71077	probable ABC transpor	3.31e+03
410	4	13.8	105	2	F71348	probable DNA-binding	3.31e+03	483	4	13.8	238	2	D71189	hypothetical protein	3.31e+03
411	4	13.8	109	4	S58648	hypothetical protein	3.31e+03	484	4	13.8	243	2	S14459	MHC H-2K1-k - mouse	3.31e+03
412	4	13.8	111	2	B70401	hypothetical protein	3.31e+03	485	4	13.8	244	2	S29982	class II histocompati	3.31e+03
413	4	13.8	113	2	S27226	NADH dehydrogenase (u	3.31e+03	486	4	13.8	245	2	S29980	class II histocompati	3.31e+03
414	4	13.8	113	3	JE0380	NADH:ubiquinone oxido	3.31e+03	487	4	13.8	246	2	I50127	MHC class II histocom	3.31e+03
415	4	13.8	114	2	JE0178	SH3 binding glutamate	3.31e+03	488	4	13.8	247	2	I51060	MHC class II beta cha	3.31e+03
416	4	13.8	120	2	B71020	hypothetical protein	3.31e+03	489	4	13.8	247	2	I51059	MHC class II beta cha	3.31e+03
417	4	13.8	124	2	I50095	MHC class II beta cha	3.31e+03	490	4	13.8	247	2	JC6540	placenta specific-hom	3.31e+03
418	4	13.8	125	2	F71156	hypothetical protein	3.31e+03	491	4	13.8	248	2	I50093	MHC class II beta cha	3.31e+03
419	4	13.8	126	2	C53116	Ig epsilon chain C re	3.31e+03	492	4	13.8	248	2	A49515	phosphoprotein 41 - h	3.31e+03
420	4	13.8	126	2	A53260	hypothetical protein	3.31e+03	493	4	13.8	249	2	I50464	MHC class II beta cha	3.31e+03
421	4	13.8	126	2	F71185	hypothetical protein	3.31e+03	494	4	13.8	252	2	I50126	MHC class II histocom	3.31e+03
422	4	13.8	127	2	PC2036	microfibril-associate	3.31e+03	495	4	13.8	254	2	F71203	hypothetical protein	3.31e+03
423	4	13.8	129	2	B71259	hypothetical protein	3.31e+03	496	4	13.8	256	2	S42932	probable transmembran	3.31e+03
424	4	13.8	133	2	I53295	bcl-2-associated prot	3.31e+03	497	4	13.8	257	2	B70452	conserved hypothetical	3.31e+03
425	4	13.8	138	2	G71065	hypothetical protein	3.31e+03	498	4	13.8	257	2	A25394	SURF-2 protein - mous	3.31e+03
426	4	13.8	138	2	S72482	hypothetical protein	3.31e+03	499	4	13.8	257	2	S29981	class II histocompati	3.31e+03
427	4	13.8	143	2	D70332	hypothetical protein	3.31e+03	500	4	13.8	259	2	D70342	hypothetical protein	3.31e+03
428	4	13.8	151	2	B71166	hypothetical protein	3.31e+03	501	4	13.8	260	2	E71045	probable ABC transpor	3.31e+03
429	4	13.8	153	2	A70456	hypothetical protein	3.31e+03	502	4	13.8	261	2	A55242	MHC class II histocom	3.31e+03
430	4	13.8	154	2	S58015	probable olfactory re	3.31e+03	503	4	13.8	261	2	S17889	class II histocompati	3.31e+03
431	4	13.8	154	2	S40522	Huntington's disease	3.31e+03	504	4	13.8	262	2	D71068	probable pseudouridyl	3.31e+03
432	4	13.8	155	2	S57426	cysteine proteinase (	3.31e+03	505	4	13.8	263	2	I37533	MHC class II histocom	3.31e+03
433	4	13.8	156	2	A70412	hypothetical protein	3.31e+03	506	4	13.8	264	2	S60367	transcription factor	3.31e+03
434	4	13.8	158	2	S71113	ubiquitin-conjugating	3.31e+03	507	4	13.8	265	2	H71299	probable hemolysin (t	3.31e+03
435	4	13.8	158	2	JC6056	ubiquitin conjugating	3.31e+03	508	4	13.8	267	2	B71082	hypothetical protein	3.31e+03
436	4	13.8	159	2	C56236	probable RNA helicase	3.31e+03	509	4	13.8	269	2	A56487	signal recognition pa	3.31e+03
437	4	13.8	162	2	F71319	hypothetical protein	3.31e+03	510	4	13.8	272	2	D70312	hypothetical protein	3.31e+03
438	4	13.8	170	2	G71323	probable purine-bindi	3.31e+03	511	4	13.8	272	2	A34693	beta-globin DNA-bindi	3.31e+03
439	4	13.8	174	2	A27293	probable Y chromosome	3.31e+03	512	4	13.8	277	2	I60122	rsu-1 homolog - human	3.31e+03
440	4	13.8	175	2	F70335	hypothetical protein	3.31e+03	513	4	13.8	278	2	S36101	CAMP response element	3.31e+03
441	4	13.8	175	2	A26882	pi12 hypothetical pro	3.31e+03	514	4	13.8	279	2	C48826	high choriolytic hanc	3.31e+03
442	4	13.8	176	2	S29031	mpv17 protein - mouse	3.31e+03	515	4	13.8	285	2	C45691	main capsid protein G	3.31e+03
443	4	13.8	176	2	S45343	glomerulocytosis pr	3.31e+03	516	4	13.8	286	2	H45691	megakaryocyte growth	3.31e+03
444	4	13.8	181	2	F70437	ATP synthase F1 delta	3.31e+03	517	4	13.8	286	2	A55530	megakaryocyte growth	3.31e+03
445	4	13.8	182	2	B70335	conserved hypothetical	3.31e+03	518	4	13.8	288	2	A33828	REX1 protein - mouse	3.31e+03
446	4	13.8	184	2	A71230	hypothetical protein	3.31e+03	519	4	13.8	291	2	G71300	probable protein-meth	3.31e+03
447	4	13.8	189	3	B45190	chitin synthase wdcHS	3.31e+03	520	4	13.8	291	2	S77611	killer toxin KPI prec	3.31e+03
448	4	13.8	192	2	D47538	bcl-2-associated prot	3.31e+03	521	4	13.8	291	2	B71087	hypothetical protein	3.31e+03
449	4	13.8	192	2	H70352	hypothetical protein	3.31e+03	522	4	13.8	292	2	B70385	hypothetical protein	3.31e+03
450	4	13.8	200	2	B71366	probable thioredoxin	3.31e+03	523	4	13.8	292	2	JC4650	histamine N-methyltra	3.31e+03
451	4	13.8	202	2	D71244	hypothetical protein	3.31e+03	524	4	13.8	292	2	G01409	histamine N-methyltra	3.31e+03
452	4	13.8	208	2	S46301	fucoxanthin chlorophy	3.31e+03	525	4	13.8	293	2	C71283	hypothetical protein	3.31e+03
453	4	13.8	211	2	I40282	outer surface protein	3.31e+03	526	4	13.8	296	2	A70427	conserved hypothetical	3.31e+03
454	4	13.8	211	2	JC5106	stromal cell-derived	3.31e+03	527	4	13.8	296	2	A71232	heme O oxygenase - Aq	3.31e+03
455	4	13.8	211	2	S58875	cdn-2 protein - human	3.31e+03	528	4	13.8	296	2	D70314	heme O oxygenase - Aq	3.31e+03
456	4	13.8	213	2	H71190	hypothetical protein	3.31e+03	529	4	13.8	296	2	B43819	immediate-early trans	3.31e+03
457	4	13.8	215	2	S34163	homeotic protein Hoxd	3.31e+03	530	4	13.8	298	2	S32426	ketohexokinase (EC 2.	3.31e+03
458	4	13.8	215	2	F71101	hypothetical protein	3.31e+03	531	4	13.8	299	2	C70349	succinyl-diaminopimel	3.31e+03
459	4	13.8	217	2	I51062	MHC class II beta cha	3.31e+03	532	4	13.8	305	2	A71247	probable L-asparagin	3.31e+03
460	4	13.8	218	2	JC4788	sodium channel protei	3.31e+03	533	4	13.8	307	4	S58112	catechol 2,3-dioxygen	3.31e+03
461	4	13.8	218	2	A55734	sodium channel, volta	3.31e+03	534	4	13.8	308	2	H71008	probable ATP-binding	3.31e+03

535	4	13.8	308	2	I46019	coatomer complex epsi	3.31e+03	608	4	13.8	425	2	A46251	adipose differentiati	3.31e+03
536	4	13.8	308	2	I48080	coatomer complex epsi	3.31e+03	609	4	13.8	425	2	E71075	probable proton gluta	3.31e+03
537	4	13.8	308	2	G71110	hypothetical protein	3.31e+03	610	4	13.8	427	2	I38909	damage-specific DNA b	3.31e+03
538	4	13.8	313	2	E45774	odorant receptor 8 -	3.31e+03	611	4	13.8	428	2	S30508	probable G protein-co	3.31e+03
539	4	13.8	319	2	I51569	UVS-2 protein - Afric	3.31e+03	612	4	13.8	429	2	I48361	gene GSG-154 protein	3.31e+03
540	4	13.8	322	2	A53715	apomucin precursor -	3.31e+03	613	4	13.8	430	2	I48142	CD19 - guinea pig (fr	3.31e+03
541	4	13.8	323	2	S55724	hypothetical protein	3.31e+03	614	4	13.8	431	2	I51189	IFAPA-400 - chicken (	3.31e+03
542	4	13.8	323	2	S55348	3',5'-cyclic-nucleoti	3.31e+03	615	4	13.8	436	2	I49714	MHC H-2K/t-w5-linked	3.31e+03
543	4	13.8	326	2	Q08585	hypothetical 36.8K pr	3.31e+03	616	4	13.8	437	2	A47104	chloride channel 64K	3.31e+03
544	4	13.8	327	4	G29876	probable H+-transport	3.31e+03	617	4	13.8	439	2	B44315	cartilage oligomeric	3.31e+03
545	4	13.8	327	4	G01037	t-Cell activation pro	3.31e+03	618	4	13.8	440	2	S71795	transcription factor	3.31e+03
546	4	13.8	330	3	JE0376	Grb-2 related adaptor	3.31e+03	619	4	13.8	443	2	I46140	p62 ras-GAP associate	3.31e+03
547	4	13.8	331	2	B70336	heterodisulfide reduc	3.31e+03	620	4	13.8	443	2	A38219	GAP-associated tyrosi	3.31e+03
548	4	13.8	332	2	A50994	glycoenin glucosyltr	3.31e+03	621	4	13.8	445	2	S32036	finger protein XFG-5.	3.31e+03
549	4	13.8	334	2	A71035	probable proteinase I	3.31e+03	622	4	13.8	448	2	I50730	yes-associated protei	3.31e+03
550	4	13.8	335	2	S53352	2,4-dienoyl-CoA reduc	3.31e+03	623	4	13.8	449	2	I39358	heterogeneous nuclea	3.31e+03
551	4	13.8	335	2	E71215	hypothetical protein	3.31e+03	624	4	13.8	450	2	A45967	vitellogenin - rainbo	3.31e+03
552	4	13.8	338	2	I38170	gene hr44 protein - h	3.31e+03	625	4	13.8	450	2	A53771	transcription factor	3.31e+03
553	4	13.8	338	2	G71128	hypothetical protein	3.31e+03	626	4	13.8	454	2	S17450	probable ligand-bindi	3.31e+03
554	4	13.8	341	2	S26886	cAMP response element	3.31e+03	627	4	13.8	454	2	A56954	yes-associated protei	3.31e+03
555	4	13.8	342	2	E70463	conserved hypothetica	3.31e+03	628	4	13.8	455	2	A70461	potassium channel pro	3.31e+03
556	4	13.8	344	2	G01628	VT4 - human (fragment	3.31e+03	629	4	13.8	458	3	T01639	ribulose biphosphate	3.31e+03
557	4	13.8	344	2	JC5602	cAMP response element	3.31e+03	630	4	13.8	459	2	S68519	tub protein, testis -	3.31e+03
558	4	13.8	344	2	JC5601	cAMP response element	3.31e+03	631	4	13.8	459	2	JH0594	vasoactive intestinal	3.31e+03
559	4	13.8	345	2	JC4025	opioid-binding cell a	3.31e+03	632	4	13.8	460	2	JC2194	vasoactive intestinal	3.31e+03
560	4	13.8	345	2	S03199	opioid-binding protei	3.31e+03	633	4	13.8	461	2	S65666	actin-binding protein	3.31e+03
561	4	13.8	346	2	JC5833	transcription factor	3.31e+03	634	4	13.8	461	2	A30222	hypothetical p15kz1 p	3.31e+03
562	4	13.8	346	2	A58583	testosterone-resistan	3.31e+03	635	4	13.8	461	2	S65665	actin-binding protein	3.31e+03
563	4	13.8	348	2	H71110	probable dehydrogenas	3.31e+03	636	4	13.8	462	3	B48326	cytb intron Ia protein	3.31e+03
564	4	13.8	350	2	JC5241	activin beta E chain	3.31e+03	637	4	13.8	462	2	I51699	gene XGF 5.1C protein	3.31e+03
565	4	13.8	351	2	A45377	transcription factor	3.31e+03	638	4	13.8	462	2	I58179	5HT3 receptor subunit	3.31e+03
566	4	13.8	353	2	JC4875	zinc-finger protein z	3.31e+03	639	4	13.8	464	2	S29754	growth hormone-releas	3.31e+03
567	4	13.8	353	2	A41558	N-syndecan - rat (fra	3.31e+03	640	4	13.8	467	2	JN0616	pituitary adenylate c	3.31e+03
568	4	13.8	354	2	S59521	mucin Sac - mouse (fr	3.31e+03	641	4	13.8	469	2	F70393	hypothetical protein	3.31e+03
569	4	13.8	358	2	I46532	protein co-factor - r	3.31e+03	642	4	13.8	469	3	T01640	ribulose biphosphate	3.31e+03
570	4	13.8	365	2	A41937	choroideremia-linked	3.31e+03	643	4	13.8	470	2	S17447	probable ligand-bindi	3.31e+03
571	4	13.8	365	2	A39481	serum response factor	3.31e+03	644	4	13.8	471	2	S15035	microsome-associated	3.31e+03
572	4	13.8	366	2	A52886	cell-surface glycopro	3.31e+03	645	4	13.8	471	2	S42631	hypothetical protein	3.31e+03
573	4	13.8	369	2	F78877	cAMP responsive elem	3.31e+03	646	4	13.8	472	2	C71172	hypothetical protein	3.31e+03
574	4	13.8	371	2	B49515	phosphoprotein 75 - h	3.31e+03	647	4	13.8	472	3	T01642	ribulose biphosphate	3.31e+03
575	4	13.8	380	2	A42832	factor VIII-associate	3.31e+03	648	4	13.8	473	2	D70405	glutamyl-tRNA synthet	3.31e+03
576	4	13.8	380	2	S22415	membrane protein TGN3	3.31e+03	649	4	13.8	474	2	A35732	protective protein pr	3.31e+03
577	4	13.8	381	2	A70454	conserved hypothetica	3.31e+03	650	4	13.8	475	2	A70320	nitrate transporter -	3.31e+03
578	4	13.8	381	2	S33449	pituitary adenylate c	3.31e+03	651	4	13.8	475	2	I78878	cAMP responsive elem	3.31e+03
579	4	13.8	383	2	A53053	dentin sialoprotein p	3.31e+03	652	4	13.8	476	2	C39481	serum response factor	3.31e+03
580	4	13.8	386	2	B71066	hypothetical protein	3.31e+03	653	4	13.8	478	2	S16867	gene H5 protein - mou	3.31e+03
581	4	13.8	387	2	B71247	hypothetical protein	3.31e+03	654	4	13.8	478	4	TVHUBD	transforming protein	3.31e+03
582	4	13.8	387	2	I38449	extracellular protein	3.31e+03	655	4	13.8	480	2	A31589	protective protein pr	3.31e+03
583	4	13.8	391	2	E71125	probable aspartate am	3.31e+03	656	4	13.8	483	2	E71133	hypothetical protein	3.31e+03
584	4	13.8	391	2	D56205	regulatory protein LB	3.31e+03	657	4	13.8	484	2	A71456	hypothetical protein	3.31e+03
585	4	13.8	391	2	S40510	nucleosome assembly p	3.31e+03	658	4	13.8	485	2	PC4427	Nck, Ash and phosphol	3.31e+03
586	4	13.8	395	2	H70385	tryptophanyl-tRNA syn	3.31e+03	659	4	13.8	485	2	S52411	ZNF165 protein - huma	3.31e+03
587	4	13.8	396	2	I58168	growth factor arg3.1	3.31e+03	660	4	13.8	486	2	B39481	serum response factor	3.31e+03
588	4	13.8	397	2	F71125	hypothetical protein	3.31e+03	661	4	13.8	489	2	E40480	Na(+)-solute symporte	3.31e+03
589	4	13.8	398	2	S17428	interleukin-1 recepto	3.31e+03	662	4	13.8	489	2	S41757	5-hydroxytryptamine-3	3.31e+03
590	4	13.8	399	2	B71140	hypothetical protein	3.31e+03	663	4	13.8	491	2	A49179	Melanoma antigen homo	3.31e+03
591	4	13.8	399	2	I48705	proteinase activated	3.31e+03	664	4	13.8	493	2	S72196	X-Pro dipeptidase (EC	3.31e+03
592	4	13.8	404	2	C70403	ferredoxin oxidoreduc	3.31e+03	665	4	13.8	493	2	E71302	conserved hypothetica	3.31e+03
593	4	13.8	405	4	A61181	homeotic protein HOX1	3.31e+03	666	4	13.8	494	2	S39063	pituitary adenyllyl cy	3.31e+03
594	4	13.8	412	2	JC4921	double C2 protein bet	3.31e+03	667	4	13.8	495	2	S39061	pituitary adenyllyl cy	3.31e+03
595	4	13.8	412	2	G02099	LYSP100-A - human	3.31e+03	668	4	13.8	495	2	S36114	pituitary adenylate c	3.31e+03
596	4	13.8	414	2	I38977	TAR DNA-binding prote	3.31e+03	669	4	13.8	495	2	JC2195	vasoactive intestinal	3.31e+03
597	4	13.8	415	2	S43484	heterogeneous nuclea	3.31e+03	670	4	13.8	496	2	A37979	cartilage matrix prot	3.31e+03
598	4	13.8	415	3	T00140	capsid protein - Stap	3.31e+03	671	4	13.8	498	2	H71279	probable fema protein	3.31e+03
599	4	13.8	415	2	S68831	CMP-N-acetylneuramina	3.31e+03	672	4	13.8	500	2	S66522	cartilage matrix prot	3.31e+03
600	4	13.8	416	2	D71297	probable flagellar ho	3.31e+03	673	4	13.8	501	2	A42030	alpha-globin transcri	3.31e+03
601	4	13.8	420	2	S65084	finger protein XFG 5-	3.31e+03	674	4	13.8	502	2	F70316	conserved hypothetica	3.31e+03
602	4	13.8	420	2	F71040	hypothetical protein	3.31e+03	675	4	13.8	502	2	C56205	transcription factor	3.31e+03
603	4	13.8	424	2	F71087	hypothetical protein	3.31e+03	676	4	13.8	502	2	B53771	transcription factor	3.31e+03
604	4	13.8	424	2	S54390	portal protein gp3 -	3.31e+03	677	4	13.8	504	2	A56205	transcription factor	3.31e+03
605	4	13.8	425	2	I49367	retinoblastoma-bindin	3.31e+03	678	4	13.8	504	2	I49257	transcription factor	3.31e+03
606	4	13.8	425	2	S36112	retinoblastoma-bindin	3.31e+03	679	4	13.8	505	2	S68518	tdub protein, brain -	3.31e+03
607	4	13.8	425	2	I39181	retinoblastoma-bindin	3.31e+03	680	4	13.8	506	2	B56201	transcription factor	3.31e+03

681	4	13.8	507	2	S25831	myocyte-specific enh	3.31e+03	754	4	13.8	639	2	B48189	sodium/phosphate cotr	3.31e+03
682	4	13.8	507	2	S52348	hypothetical protein	3.31e+03	755	4	13.8	639	2	G02919	transcription factor	3.31e+03
683	4	13.8	508	2	A45477	cAMP response element	3.31e+03	756	4	13.8	642	2	C71137	hypothetical protein	3.31e+03
684	4	13.8	509	2	A48048	egg envelope protein	3.31e+03	757	4	13.8	642	2	T46534	renal sodium-dependen	3.31e+03
685	4	13.8	513	2	S47631	pituitary adenylate c	3.31e+03	758	4	13.8	644	2	A53184	myc far upstream elem	3.31e+03
686	4	13.8	514	2	B71252	probable tpr protein	3.31e+03	759	4	13.8	647	2	JC4239	phospholipase A2-acti	3.31e+03
687	4	13.8	514	2	A56201	transcription factor	3.31e+03	760	4	13.8	647	2	A57352	GPI-anchored protein	3.31e+03
688	4	13.8	514	2	A35658	transcription factor	3.31e+03	761	4	13.8	650	2	A40686	Rab geranylgeranyl tr	3.31e+03
689	4	13.8	515	2	I50637	dynein light chain A	3.31e+03	762	4	13.8	652	2	A37000	leukemia virus recept	3.31e+03
690	4	13.8	518	2	B48088	beta-transducin repea	3.31e+03	763	4	13.8	653	2	S33978	hypothetical protein	3.31e+03
691	4	13.8	518	2	G70471	flagellin - Aquifex a	3.31e+03	764	4	13.8	653	2	S33978	choroideremia, Rab ger	3.31e+03
692	4	13.8	519	2	S45345	TUPI-like enhancer -	3.31e+03	765	4	13.8	653	2	A49722	endoglin precursor -	3.31e+03
693	4	13.8	519	2	S66673	disulfide isomerase-r	3.31e+03	766	4	13.8	653	2	A54366	sodium/phosphate cotr	3.31e+03
694	4	13.8	519	2	G78089	G-protein signalling	3.31e+03	767	4	13.8	656	2	A49579	amphotropic murine re	3.31e+03
695	4	13.8	523	2	S39062	pituitary adenylate cy	3.31e+03	768	4	13.8	659	2	A35723	70K peroxisomal membr	3.31e+03
696	4	13.8	523	2	S39062	pituitary adenylate c	3.31e+03	769	4	13.8	660	2	G71251	regulatory protein Ne	3.31e+03
697	4	13.8	526	2	A34855	67K neutrophil oxidas	3.31e+03	770	4	13.8	660	2	H71376	probable tpr protein	3.31e+03
698	4	13.8	533	2	D71338	probable ribose/galac	3.31e+03	771	4	13.8	660	2	S71949	metalloproteinase 12	3.31e+03
699	4	13.8	538	2	E71255	conserved hypotheticala	3.31e+03	772	4	13.8	661	2	A53668	glycoprotein gp100 pr	3.31e+03
700	4	13.8	540	2	A70358	topoisomerase I - Aqu	3.31e+03	773	4	13.8	661	2	A55877	tumor necrosis factor	3.31e+03
701	4	13.8	540	2	JC4916	signal transducing ad	3.31e+03	774	4	13.8	662	2	I38400	melanoma-associated M	3.31e+03
702	4	13.8	541	2	JC5423	2-hydroxyacylsphingos	3.31e+03	775	4	13.8	665	2	G71251	probable membrane pro	3.31e+03
703	4	13.8	542	2	B56205	transcription factor	3.31e+03	776	4	13.8	668	2	A41234	melanocyte-specific p	3.31e+03
704	4	13.8	545	2	I48951	oncofetal antigen - m	3.31e+03	777	4	13.8	671	2	B54857	transcription factor	3.31e+03
705	4	13.8	547	2	B45808	B-lymphocyte antigen	3.31e+03	778	4	13.8	677	2	I49045	SKD3 - mouse	3.31e+03
706	4	13.8	547	2	I49580	CD19 - mouse	3.31e+03	779	4	13.8	679	2	E71282	probable NH(3)-depend	3.31e+03
707	4	13.8	548	2	I68667	transcription factor	3.31e+03	780	4	13.8	681	2	I78558	hypothetical Brachyur	3.31e+03
708	4	13.8	548	2	S59133	ETS2 repressor factor	3.31e+03	781	4	13.8	684	2	S60266	novel antigen recepto	3.31e+03
709	4	13.8	548	2	JC4917	signal transducing ad	3.31e+03	782	4	13.8	686	2	JC5708	villin-like protein -	3.31e+03
710	4	13.8	548	2	S52735	CW17R protein - mouse	3.31e+03	783	4	13.8	697	2	B44489	G box-binding protei	3.31e+03
711	4	13.8	552	2	S71811	probable transcriptio	3.31e+03	784	4	13.8	698	2	A57644	Na+/H+-exchanging pro	3.31e+03
712	4	13.8	556	2	A44441	B-cell antigen CD19 p	3.31e+03	785	4	13.8	701	2	I56208	heat shock protein 70	3.31e+03
713	4	13.8	559	2	B56731	chromatin assembly fa	3.31e+03	786	4	13.8	704	2	A48097	rabphilin-3A - bovine	3.31e+03
714	4	13.8	563	2	A41530	src substrate p80/85	3.31e+03	787	4	13.8	706	2	H70425	conserved hypotheticala	3.31e+03
715	4	13.8	565	2	JE0160	sodium bicarbonate co	3.31e+03	788	4	13.8	706	2	G71004	hypothetical protein	3.31e+03
716	4	13.8	568	2	I58106	gene DMR-N9 protein -	3.31e+03	789	4	13.8	720	2	S69890	mitogen inducible gen	3.31e+03
717	4	13.8	570	2	I59402	NAB1 - rat	3.31e+03	790	4	13.8	720	2	G70320	aldehyde dehydrogenas	3.31e+03
718	4	13.8	570	2	A57535	intraleukin 1 receptor	3.31e+03	791	4	13.8	726	2	S18208	adducin beta - human	3.31e+03
719	4	13.8	573	2	C71170	hypothetical protein	3.31e+03	792	4	13.8	729	2	S68191	triadin - human	3.31e+03
720	4	13.8	574	2	A46054	GTP-binding protein A	3.31e+03	793	4	13.8	729	2	I54501	interferon gamma-indu	3.31e+03
721	4	13.8	574	2	S48860	gene s17 protein - mo	3.31e+03	794	4	13.8	735	2	JC5869	beta-glucosidase (EC	3.31e+03
722	4	13.8	575	2	I59327	olfactory cyclic nucl f	3.31e+03	795	4	13.8	736	2	I51691	dishevelled homolog -	3.31e+03
723	4	13.8	576	2	A48157	renal transcription fac	3.31e+03	796	4	13.8	741	2	JC2074	alpha-1,3(6)-mannosyl	3.31e+03
724	4	13.8	581	2	G02318	Ets transcription fac	3.31e+03	797	4	13.8	742	2	S12533	zfa protein - mouse	3.31e+03
725	4	13.8	582	2	H70426	neutral proteinase (E	3.31e+03	798	4	13.8	743	2	F71062	hypothetical protein	3.31e+03
726	4	13.8	582	2	I50224	ECH - chicken	3.31e+03	799	4	13.8	743	2	T00634	hypothetical protein	3.31e+03
727	4	13.8	585	2	A46507	Ig alpha chain - chic	3.31e+03	800	4	13.8	753	2	S48059	metal-regulatory tran	3.31e+03
728	4	13.8	586	2	S22156	transforming protein	3.31e+03	801	4	13.8	757	2	A39283	gamma-glutamyl carbox	3.31e+03
729	4	13.8	589	2	I38598	zinc finger protein z	3.31e+03	802	4	13.8	763	2	JC4376	beta-glucosidase (EC	3.31e+03
730	4	13.8	590	2	A36535	gadd34 protein - long	3.31e+03	803	4	13.8	763	2	A43314	special AT-rich seque	3.31e+03
731	4	13.8	596	2	I38228	Shb protein - human	3.31e+03	804	4	13.8	764	2	A47456	down-regulated in ade	3.31e+03
732	4	13.8	597	2	A30814	development-specific	3.31e+03	805	4	13.8	766	2	S45344	TUPI like enhancer -	3.31e+03
733	4	13.8	599	2	S38735	centromere protein B	3.31e+03	806	4	13.8	767	2	S47639	acetylcholinesterase	3.31e+03
734	4	13.8	600	2	S56744	mucin (clone pgm7-1)	3.31e+03	807	4	13.8	770	2	G02228	DOC-2 - human	3.31e+03
735	4	13.8	606	2	S70358	centromere protein B	3.31e+03	808	4	13.8	771	2	B38252	granulocyte colony-st	3.31e+03
736	4	13.8	609	2	S55270	catrictollastatin prec	3.31e+03	809	4	13.8	775	2	I49759	Hrs - mouse	3.31e+03
737	4	13.8	613	2	A42021	artificial-warming sequ	3.31e+03	810	4	13.8	776	2	I48317	cyclin F - mouse	3.31e+03
738	4	13.8	614	2	C71320	probable adenylate cy	3.31e+03	811	4	13.8	779	2	S36111	osteoblast-specific f	3.31e+03
739	4	13.8	616	2	A5796	ecarin precursor - sa	3.31e+03	812	4	13.8	783	2	JH0329	granulocyte colony-st	3.31e+03
740	4	13.8	623	2	I54371	transcription factor	3.31e+03	813	4	13.8	784	2	A33633	transcription repress	3.31e+03
741	4	13.8	623	2	A49112	sodium-glucose cotran	3.31e+03	814	4	13.8	786	2	S22155	oncogene 1 (tre-2 loc	3.31e+03
742	4	13.8	626	2	I49100	musc6 precursor - mou	3.31e+03	815	4	13.8	786	2	I49274	mammary gland factor	3.31e+03
743	4	13.8	626	2	A42891	beta-galactosidase (E	3.31e+03	816	4	13.8	786	2	S54773	stat5b protein - mous	3.31e+03
744	4	13.8	626	2	S38871	Pmel 17 protein - mou	3.31e+03	817	4	13.8	786	2	A55501	cyclin F - human	3.31e+03
745	4	13.8	627	2	A40201	artificial-warming sequ	3.31e+03	818	4	13.8	793	2	S54772	mammary gland factor	3.31e+03
746	4	13.8	629	2	A71023	probable arginyl-tRNA	3.31e+03	819	4	13.8	794	2	G02317	transcription activat	3.31e+03
747	4	13.8	630	2	T00351	hypothetical protein	3.31e+03	820	4	13.8	794	2	S59069	z13 protein - mouse	3.31e+03
748	4	13.8	633	2	A36353	DNA repair protein XR	3.31e+03	821	4	13.8	794	2	S55527	mammary gland factor	3.31e+03
749	4	13.8	634	2	T00359	hypothetical protein	3.31e+03	822	4	13.8	796	2	A45695	capsid precursor - hu	3.31e+03
750	4	13.8	637	2	A54000	sodium/phosphate cotr	3.31e+03	823	4	13.8	800	2	I51653	dsRNA-binding protein	3.31e+03
751	4	13.8	637	2	A48189	sodium/phosphate cotr	3.31e+03	824	4	13.8	809	2	A34404	oxysterol-binding pro	3.31e+03
752	4	13.8	637	2	A35015	glutamate--cysteine 1	3.31e+03	825	4	13.8	809	2	A46747	Na+/H+-exchanging pro	3.31e+03
753	4	13.8	637	2	JH0611	glutamate--cysteine 1	3.31e+03	826	4	13.8	811	2	S36109	osteoblast-specific f	3.31e+03

827	4	13.8	813	2	A46748	Na+/H+-exchanging pro	3.31e+03	900	4	13.8	1154	2	A56242	E-box-binding repres	3.31e+03
828	4	13.8	815	3	T00264	high carbon dioxide r	3.31e+03	901	4	13.8	1168	2	I56985	kalinin B1 - mouse	3.31e+03
829	4	13.8	831	2	JQ1655	prolactin receptor pr	3.31e+03	902	4	13.8	1170	2	A53612	laminin B1k chain pre	3.31e+03
830	4	13.8	834	2	S66258	glucosidase I - huma	3.31e+03	903	4	13.8	1175	2	S52417	E-selectin ligand-1 p	3.31e+03
831	4	13.8	836	2	S36110	osteoblast-specific f	3.31e+03	904	4	13.8	1184	2	T00253	gene Ankhzn protein -	3.31e+03
832	4	13.8	837	2	I57557	DNA-Binding Protein a	3.31e+03	905	4	13.8	1184	2	I49017	calcium-activated pot	3.31e+03
833	4	13.8	837	2	A42112	mucin-like peptide ML	3.31e+03	906	4	13.8	1190	2	S21977	pm5 protein - human	3.31e+03
834	4	13.8	840	2	A48998	granulocyte colony-st	3.31e+03	907	4	13.8	1196	2	A48206	calcium-activated pot	3.31e+03
835	4	13.8	847	2	I48968	precursor of major fi	3.31e+03	908	4	13.8	1196	2	S35994	DNA repair protein xp	3.31e+03
836	4	13.8	848	2	A54740	interleukin-4-induced	3.31e+03	909	4	13.8	1199	2	A41939	G protein-coupled glu	3.31e+03
837	4	13.8	849	2	I49060	major 80,000 Mr fibro	3.31e+03	910	4	13.8	1199	2	S20969	Na+/Ca2+,K+-exchangi	3.31e+03
838	4	13.8	850	2	S56015	gastric mucin - huma	3.31e+03	911	4	13.8	1207	2	S27954	leucine-rich protein	3.31e+03
839	4	13.8	852	2	I49583	differentiation antig	3.31e+03	912	4	13.8	1216	2	JW0105	synaptotagmin 2 alpa	3.31e+03
840	4	13.8	863	2	C38252	granulocyte colony-st	3.31e+03	913	4	13.8	1217	2	T00270	hypothetical protein	3.31e+03
841	4	13.8	866	2	JC4305	dynamitin II - huma	3.31e+03	914	4	13.8	1234	2	T00363	hypothetical protein	3.31e+03
842	4	13.8	868	2	A46512	CD22 homolog/B lympho	3.31e+03	915	4	13.8	1252	2	T00263	hypothetical protein	3.31e+03
843	4	13.8	868	2	A36878	dynamitin 2 - rat	3.31e+03	916	4	13.8	1256	2	JEO209	brain-specific angio	3.31e+03
844	4	13.8	869	2	A25945	coagulation factor VI	3.31e+03	917	4	13.8	1265	2	S57968	Ran-binding protein 2	3.31e+03
845	4	13.8	870	2	A53165	dynamitin II isoform aa	3.31e+03	918	4	13.8	1273	2	T00338	hypothetical protein	3.31e+03
846	4	13.8	871	2	B53165	dynamitin II isoform ba	3.31e+03	919	4	13.8	1274	2	S55050	cardiac myosin-bindin	3.31e+03
847	4	13.8	872	2	B71039	hypothetical protein	3.31e+03	920	4	13.8	1278	3	A47462	probable DNA-directed	3.31e+03
848	4	13.8	875	2	I59350	karyopherin beta - ra	3.31e+03	921	4	13.8	1280	2	T00365	hypothetical protein	3.31e+03
849	4	13.8	876	2	I52907	importin beta chain -	3.31e+03	922	4	13.8	1281	2	T00346	hypothetical protein	3.31e+03
850	4	13.8	876	2	S56288	nuclear pore-targetin	3.31e+03	923	4	13.8	1287	2	A41685	SIL protein - human	3.31e+03
851	4	13.8	882	2	A53736	mastocytoma glucosam	3.31e+03	924	4	13.8	1289	2	T00387	KIAA0622 protein - hu	3.31e+03
852	4	13.8	883	2	A49733	heparan sulfate-N-dea	3.31e+03	925	4	13.8	1291	2	H71143	probable ribonucleosi	3.31e+03
853	4	13.8	886	2	S36326	clathrin assembly pro	3.31e+03	926	4	13.8	1299	2	T00261	hypothetical protein	3.31e+03
854	4	13.8	889	2	G71453	hypothetical protein	3.31e+03	927	4	13.8	1300	2	I53799	CGI protein - human	3.31e+03
855	4	13.8	901	2	A44825	phosphoprotein, synap	3.31e+03	928	4	13.8	1317	2	T03748	apoptosis associated	3.31e+03
856	4	13.8	901	3	T03726	capsid polypeptide -	3.31e+03	929	4	13.8	1320	2	JC5630	TCOP1 protein - mouse	3.31e+03
857	4	13.8	902	3	T01668	pol polypeptide - hu	3.31e+03	930	4	13.8	1324	2	T00386	hypothetical protein	3.31e+03
858	4	13.8	906	2	T00039	hypothetical protein	3.31e+03	931	4	13.8	1356	2	S32763	kinectin 1 - human	3.31e+03
859	4	13.8	910	2	S68983	auxilin - bovine	3.31e+03	932	4	13.8	1361	3	T03415	S-layer protein - Cam	3.31e+03
860	4	13.8	915	2	S36327	clathrin assembly pro	3.31e+03	933	4	13.8	1373	2	JEO095	gastric mucin MUC5AC	3.31e+03
861	4	13.8	925	2	T00334	hypothetical protein	3.31e+03	934	4	13.8	1375	2	F48216	neurexin III-alpha se	3.31e+03
862	4	13.8	940	2	T00056	hypothetical protein	3.31e+03	935	4	13.8	1378	2	E48216	neurexin III-alpha se	3.31e+03
863	4	13.8	941	2	F71332	probable chromosome s	3.31e+03	936	4	13.8	1384	3	T02748	hypothetical protein	3.31e+03
864	4	13.8	943	2	A42881	centromere protein C	3.31e+03	937	4	13.8	1386	2	T00257	hypothetical protein	3.31e+03
865	4	13.8	946	3	T00024	ent-Kaurene synthase	3.31e+03	938	4	13.8	1388	2	T00063	hypothetical protein	3.31e+03
866	4	13.8	947	2	T00340	hypothetical protein	3.31e+03	939	4	13.8	1392	2	D48216	neurexin III-alpha se	3.31e+03
867	4	13.8	954	2	I61714	CO-repressor protein	3.31e+03	940	4	13.8	1395	2	C48216	neurexin III-alpha se	3.31e+03
868	4	13.8	963	2	A55926	DNA-binding protein R	3.31e+03	941	4	13.8	1396	2	D48216	neurexin III-alpha se	3.31e+03
869	4	13.8	975	2	I59422	rsec8 - rat (fragment	3.31e+03	942	4	13.8	1402	2	I46707	initiation factor 4-g	3.31e+03
870	4	13.8	979	2	A35913	regulatory factor X -	3.31e+03	943	4	13.8	1417	2	T03449	receptor-associated c	3.31e+03
871	4	13.8	980	2	S71090	peroxisome biogenesis	3.31e+03	944	4	13.8	1420	2	T03749	probable nuclear rece	3.31e+03
872	4	13.8	1007	2	T01437	hypothetical protein	3.31e+03	945	4	13.8	1426	2	JN0337	nitric-oxide synthase	3.31e+03
873	4	13.8	1009	2	A56095	HIP116 protein - huma	3.31e+03	946	4	13.8	1429	2	JN0609	hypothetical protein	3.31e+03
874	4	13.8	1009	2	S49618	helicase-like transcr	3.31e+03	947	4	13.8	1435	2	B48216	neurexin III-alpha se	3.31e+03
875	4	13.8	1012	2	I53172	RAE-28 - mouse	3.31e+03	948	4	13.8	1438	2	B48216	neurexin III-alpha se	3.31e+03
876	4	13.8	1016	2	T00375	KIAA0647 protein - hu	3.31e+03	949	4	13.8	1441	2	T00335	hypothetical protein	3.31e+03
877	4	13.8	1017	2	I37465	HIRA protein - huma	3.31e+03	950	4	13.8	1443	2	I50600	neogenin - chicken (f	3.31e+03
878	4	13.8	1018	2	JC5799	alpha-D-mannosidase (	3.31e+03	951	4	13.8	1447	2	A54100	tumor suppressor prot	3.31e+03
879	4	13.8	1023	2	JC4013	major acidic nuclear	3.31e+03	952	4	13.8	1468	2	C48218	neurexin III-alpha me	3.31e+03
880	4	13.8	1023	2	A47296	thiazide-sensitive el	3.31e+03	953	4	13.8	1468	2	A44345	nucleoporin - rat	3.31e+03
881	4	13.8	1042	2	A31591	transcription regulat	3.31e+03	954	4	13.8	1471	2	B48218	neurexin III-alpha me	3.31e+03
882	4	13.8	1042	2	S41705	EVII protein - huma	3.31e+03	955	4	13.8	1472	2	B54774	ATP binding cassette	3.31e+03
883	4	13.8	1042	2	S42511	RAG-1 protein - rabbi	3.31e+03	956	4	13.8	1475	2	S42718	nuclear pore complex	3.31e+03
884	4	13.8	1045	2	JC5795	CDEP protein - huma	3.31e+03	957	4	13.8	1553	2	T03301	rab3 effector protein	3.31e+03
885	4	13.8	1051	2	A60191	oncogene Evi-1 - huma	3.31e+03	958	4	13.8	1560	2	I54361	escapes X-chromosome	3.31e+03
886	4	13.8	1052	2	T00067	hypothetical protein	3.31e+03	959	4	13.8	1571	2	T00062	hypothetical protein	3.31e+03
887	4	13.8	1059	2	T00377	KIAA0642 protein - hu	3.31e+03	960	4	13.8	1572	2	T00027	brain-specific angio	3.31e+03
888	4	13.8	1072	2	A37127	microtubule-associate	3.31e+03	961	4	13.8	1575	2	A48218	neurexin III-alpha me	3.31e+03
889	4	13.8	1081	2	T00330	hypothetical protein	3.31e+03	962	4	13.8	1578	2	J48216	neurexin III-alpha me	3.31e+03
890	4	13.8	1083	2	S54293	regulator protein p12	3.31e+03	963	4	13.8	1613	2	JEO272	low density lipoprote	3.31e+03
891	4	13.8	1089	2	S22158	transforming protein	3.31e+03	964	4	13.8	1621	2	S62356	TRP-185 protein - hum	3.31e+03
892	4	13.8	1091	2	E71322	probable isoleucyl-tr	3.31e+03	965	4	13.8	1651	2	JC1340	outer membrane protei	3.31e+03
893	4	13.8	1114	3	JH0284	125K surface antigen	3.31e+03	966	4	13.8	1659	2	JC4956	vitellogenin precurs	3.31e+03
894	4	13.8	1123	2	S36846	myosin-binding protein	3.31e+03	967	4	13.8	1704	2	S71363	probable transport pr	3.31e+03
895	4	13.8	1124	2	JX0293	zinc finger protein A	3.31e+03	968	4	13.8	1733	2	S27939	tensin - chicken	3.31e+03
896	4	13.8	1125	2	B41206	microtubule-associate	3.31e+03	969	4	13.8	1737	2	T00209	MEGF8 protein - huma	3.31e+03
897	4	13.8	1138	2	S24614	myosin-binding protei	3.31e+03	970	4	13.8	1737	2	A37491	hypothetical helicase	3.31e+03
898	4	13.8	1142	2	T00022	Bl20 protein - huma	3.31e+03	971	4	13.8	1744	2	A54970	tensin, cardiac muscl	3.31e+03
899	4	13.8	1152	2	A33183	microtubule-associate	3.31e+03	972	4	13.8	1745	2	A46431	tight junction-associ	3.31e+03



```

973      4 13.8 1753 2 T00350      hypothetical protein      3.31e+03
974      4 13.8 1759 3 T03725      replicase polyprotein  3.31e+03
975      4 13.8 1792 2 A57075      tensin - chicken (fra 3.31e+03
976      4 13.8 1820 2 S71853      genome polyprotein -  3.31e+03
977      4 13.8 1820 2 A55494      latent transforming g  3.31e+03
978      4 13.8 1886 2 S04921      nuclear pore protein  3.31e+03
979      4 13.8 1894 2 JC4980      plexin 1 precursor -  3.31e+03
980      4 13.8 1905 2 I51553      Plexin - African claw 3.31e+03
981      4 13.8 1920 2 A53188      pericentrin - mouse  3.31e+03
982      4 13.8 1922 2 T00637      hypothetical protein  3.31e+03
983      4 13.8 2055 2 T00093      hypothetical protein  3.31e+03
984      4 13.8 2139 2 A44467      voltage-dependent cal 3.31e+03
985      4 13.8 2140 2 JH0426      voltage-dependent cal 3.31e+03
986      4 13.8 2143 2 JH0427      voltage-dependent cal 3.31e+03
987      4 13.8 2223 2 A74447      calcium channel prote 3.31e+03
988      4 13.8 2871 2 A55567      fibrillin 1 - bovine  3.31e+03
989      4 13.8 2871 2 A55624      fibrillin-1 precursor 3.31e+03
990      4 13.8 2907 2 A57278      fibrillin-2 precursor 3.31e+03
991      4 13.8 3002 2 A47221      fibrillin 1 precursor 3.31e+03
992      4 13.8 3119 2 I49729      HD protein - mouse    3.31e+03
993      4 13.8 3144 2 A45068      Huntington disease-as 3.31e+03
994      4 13.8 3224 2 S58884      Ran-binding protein 2  3.31e+03
995      4 13.8 3707 2 S18252      heparan sulfate prote 3.31e+03
996      4 13.8 4436 2 E17086      hypothetical protein   3.31e+03
997      4 13.8 4861 2 S71752      giant protein p619 -  3.31e+03
998      4 13.8 4957 2 T03455      ALR protein - human   3.31e+03
999      4 13.8 5262 2 T03454      ALR protein - human   3.31e+03
1000     4 13.8 13288 2 T03099      apomucin, submaxillar 3.31e+03

ALIGNMENTS

RESULT 1
ENTRY 1JHU3      #type complete
TITLE Desmoglein 3 precursor - human
ALTERNATE_NAMES pemphigus vulgaris antigen
ORGANISM #formal_name Homo sapiens #common_name man
DATE 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change
18-Sep-1998
ACCESSIONS A41088
REFERENCE A41088
#authors Amagai, M.; Klaus-Kovtun, V.; Stanley, J.R.
#journal Cell (1991) 67:869-877
#title Autoantibodies against a novel epithelial cadherin in
pemphigus vulgaris, a disease of cell adhesion.
#cross-references MUID:92069753
#accession A41088
#molecule_type mRNA
#residues 1-999 #label AMA
#cross-references GB:M76482; NID:g190751; PID:g190752

GENETICS
#gene GDB:DSG3
#map_position 18q12.1-18q12.2
CLASSIFICATION #superfamily cadherin; cadherin repeat homology
KEYWORDS calcium binding; cell adhesion; duplication; glycoprotein;
transmembrane protein

FEATURE
1-23 #domain signal sequence #status predicted #label SIG\
24-49 #domain propeptide #status predicted #label PRO\
50-999 #product desmoglein homolog #status predicted #label
NAT\
50-615 #domain extracellular #status predicted #label EXT\
52-157 #domain cadherin repeat homology #label CR1\
160-267 #domain cadherin repeat homology #label CR2\
270-383 #domain cadherin repeat homology #label CR3\
390-495 #domain cadherin repeat homology #label CR4\
496-598 #domain cadherin repeat homology #label CR5\
616-639 #domain transmembrane #status predicted #label TM\
640-999 #domain intracellular #status predicted #label INT\
910-938 #domain desmoglein repeat #label DG1\
937-966 #domain desmoglein repeat #label DG2\
110,180,545 #binding_site carbohydrate (Asn) (covalent) #status
```

```

SUMMARY #length 999 #molecular-weight 107502 #checksum 8311
predicted
Query Match 24.1%; Score 7; DB 1; Length 999;
Best Local Similarity 100.0%; Pred. No. 5.71e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 100 FVVDKNT 106
QY 22 FVVDKNT 28
|||||

RESULT 2
ENTRY JC4849 #type complete
TITLE H+-transporting ATPase (EC 3.6.1.35) 14K chain, vacuolar -
fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change
17-Mar-1999
ACCESSIONS JC4849; S38436
REFERENCE JC4849
#authors Guo, Y.; Kaiser, K.; Wieczorek, H.; Dow, A.T.
#journal Gene (1996) 172:239-243
#title The Drosophila melanogaster gene vha14 encoding a 14-kDa
F-subunit of the vacuolar ATPase.
#cross-references MUID:96269411
#accession JC4849
#molecule_type mRNA
#residues 1-124 #label GUO
#cross-references EMBL:Z26918; NID:g408122; PID:g408123
#note submitted to the EMBL Data Library, October 1993
COMMENT This enzyme is an ubiquitous proton pump. It pumps protons from the
cytoplasm to the internal space of organelles, and plays a role
in energising plasma membrane ion transport.

GENETICS
#gene FlyBase:Vha14
#cross-references FlyBase:FBgn0010426
KEYWORDS hydrolase
SUMMARY #length 124 #molecular-weight 13850 #checksum 3615

Query Match 20.7%; Score 6; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 VVDKNT 47
QY 23 VVDKNT 28
|||||

RESULT 3
ENTRY A53055 #type complete
TITLE H+-transporting ATPase (EC 3.6.1.35) 14K chain, vacuolar -
tobacco hornworm
ORGANISM #formal_name Manduca sexta #common_name tobacco hornworm
DATE 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change
09-Sep-1997
ACCESSIONS A53055; S25013
REFERENCE A53055
#authors Graef, R.; Lepier, A.; Harvey, W.R.; Wieczorek, H.
#journal J. Biol. Chem. (1994) 269:3767-3774
#title A novel 14-kDa V-ATPase subunit in the tobacco hornworm
midgut.
#accession A53055
#molecule_type mRNA
#residues 1-124 #label GRA
#cross-references EMBL:X67130; NID:g9728; PID:g9729
KEYWORDS hydrolase
SUMMARY #length 124 #molecular-weight 13814 #checksum 4156

Query Match 20.7%; Score 6; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

Db      42 VVDKNT 47
QY      23 VVDKNT 28

RESULT  4
ENTRY   S48417      #type complete
TITLE   hypothetical protein YIL062c - yeast (Saccharomyces
         cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE     02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change
         06-Feb-1998
ACCESSIONS S48417
REFERENCE   S48407
#authors   Smith, V.
#submission submitted to the EMBL Data Library, September 1994
#accession S48417
#molecule_type DNA
#residues  1-154 #label SMI
#cross-references GB:247047; EMBL:238060; NID:g603997; PID:g763284;
MIPS:YIL062c

GENETICS
#gene      SGD:ARC15
#map_position 9L
#cross-references SGD:S0001324; MIPS:YIL062c
SUMMARY   #length 154 #molecular-weight 17134 #checksum 1071

Query Match      20.7%; Score 6; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      121 GQKQGG 126
QY      16 GQKQGG 21

RESULT  5
ENTRY   S68597      #type complete
TITLE   NADPH dehydrogenase (flavin) (EC 1.6.8.2) - bovine
ALTERNATE_NAMES diaphorase; flavin reductase; green heme-binding protein;
                 methemoglobin reductase
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE     24-Aug-1996 #sequence_revision 25-Apr-1997 #text_change
         30-Jan-1998
ACCESSIONS S68597; S77619; S68275; S69235; S69236
REFERENCE   A58191
#authors   Quandt, K.S.; Hultquist, D.E.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1994) 91:9322-9326
#title     Flavin reductase: sequence of cDNA from bovine liver and
           tissue distribution.
#cross-references MUID:95023903
#accession S68597
#molecule_type mRNA
#residues  1-206 #label OUA
#cross-references EMBL:L35045; NID:g516593; PID:g516594
#experimental_source liver
#accession S77619
#molecule_type protein
#residues  2-58;118-169 #label QUX
REFERENCE   S68275
#authors   Shalloe, F.; Elliott, G.; Ennis, O.; Mantle, T.J.
#journal   Biochem. J. (1996) 316:385-387
#title     Evidence that biliverdin-IX-beta reductase and flavin
           reductase are identical.
#accession S68275
#molecule_type protein
#residues  2-14,'Vn' #label SHA
#experimental_source erythrocytes
#note      5-Leu, 16-Gly and 16-Val were also found
REFERENCE   S69235
#authors   Quandt, K.S.; Xu, F.; Chen, P.; Hultquist, D.E.
#journal   Biochem. Biophys. Res. Commun. (1991) 178:315-321
#title     Evidence that the protein components of bovine erythrocyte

```

---

```

green heme binding protein and flavin reductase are
identical.
#cross-references MUID:91298967
#accession S69235
#molecule_type protein
#residues  2-18,'X',20-30,'X',32,'X',34-37;121-122,'X',124-137;
         146-151,'X',153-154,'E',156-169 #label QUM
#note      isolated as green heme-binding protein
#accession S69236
#molecule_type protein
#residues  2-33,'X',35-37 #label QUF
#note      isolated as flavin reductase
KEYWORDS   NADP; oxidoreductase
FEATURE    2-206
SUMMARY   #product NADPH dehydrogenase (flavin) #status
           experimental #label MAT
           #length 206 #molecular-weight 22132 #checksum 326

Query Match      20.7%; Score 6; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      114 LLWDPS 119
QY      10 LLWDPS 15

RESULT  6
ENTRY   A71708      #type complete
TITLE   acyl-[acyl-carrier-protein]--UDP-n-acetylglucosamine
           O-acyltransferase (lpxA) RP007 - Rickettsia prowazekii
ORGANISM #formal_name Rickettsia prowazekii
DATE     21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change
         21-Nov-1998
ACCESSIONS A71708
REFERENCE   A71630
#authors   Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.;
           Sacheritz-Ponten, T.; Alsmark, U.C.M.; Podowski, R.M.;
           Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurland,
           C.G.
#journal   Nature (1998) 396:133-140
#title     The genome sequence of Rickettsia prowazekii and the origin
           of mitochondria.
#accession A71708
#status    preliminary; nucleic acid sequence not shown;
           translation not shown
#molecule_type DNA
#residues  1-264 #label AND
#cross-references GB:AJ235270; GB:AJ235269; NID:g3860572; PID:el342323;
           PID:g3860579
#experimental_source strain Madrid E

GENETICS
#gene      lpxA; RP007
#length 264 #molecular-weight 28470 #checksum 2368
SUMMARY

Query Match      20.7%; Score 6; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      68 QPPQIL 73
QY      5 QPPQIL 10

RESULT  7
ENTRY   JE0174      #type complete
TITLE   frizzled protein-2 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE     03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change
         10-Jul-1998
ACCESSIONS JE0174
REFERENCE   JE0174
#authors   Hu, E.; Zhu, Y.; Fredrickson, T.; Barnes, M.; Kelsell, D.;

```

```

Beeley, L.; Brooks, D.
Biochem. Biophys. Res. Commun. (1998) 247:287-293
#journal
#title Tissue restricted expression of two human frzbs in
        preadipocytes and pancreas.
#accession
#molecule_type mRNA
#residues 1-295 #label HUA
GENETICS
#map_position 4q
SUMMARY
#length 295 #molecular-weight 33490 #checksum 279
Query Match 20.7%; Score 6; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 GQKGG 266
Qy 16 GQKGG 21

RESULT 8
ENTRY
#molecule_type DNA
#type complete
TITLE
hypothetical protein PH1228 - Pyrococcus horikoshii
ORGANISM
#formal_name Pyrococcus horikoshii
DATE 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change
05-Dec-1998
ACCESSIONS
REFERENCE
#authors
Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.;
Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi,
A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.
DNA Res. (1998) 5:55-76
#journal
#title Complete sequence and gene organization of the genome of a
        hyperthermophilic archaeobacterium, Pyrococcus horikoshii
        OT3.
#cross-references MUID:98344137
#accession
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-325 #label KAW
#cross-references GB:AP000005; NID:g3236132; PID:d1031271; PID:g3257645
#experimental_source strain OT3
#note this accession replaces an interim accession for a
        sequence replaced by GenBank

GENETICS
#gene
PH1228
CLASSIFICATION
#superfamily conserved hypothetical protein MJ1427
SUMMARY
#length 325 #molecular-weight 36135 #checksum 141
Query Match 20.7%; Score 6; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 138 GGFVVD 143
Qy 20 GGFVVD 25

RESULT 9
ENTRY
#molecule_type DNA
#type complete
TITLE
flagellar switch protein FlgC - Aquifex aeolicus
ORGANISM
#formal_name Aquifex aeolicus
DATE 08-May-1998 #sequence_revision 08-May-1998 #text_change
21-Aug-1998
ACCESSIONS
REFERENCE
#authors
Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;
Lenox, A.L.; Graham, D.E.; Overbeek, R.; Sneed, M.A.;

```

```

Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short,
J.M.; Olson, G.J.; Swanson, R.V.
Nature (1998) 392:353-358
#journal
#title The complete genome of the hyperthermophilic bacterium
        Aquifex aeolicus.
#cross-references MUID:98196666
#accession
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-328 #label AQF
#cross-references GB:AE000700; NID:g2983248; PID:g2983250; GB:AE000657
#experimental_source strain VF5
GENETICS
#gene
flgC
CLASSIFICATION
#superfamily flagellar switch protein flgC
SUMMARY
#length 328 #molecular-weight 37032 #checksum 8398
Query Match 20.7%; Score 6; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 256 VDKNT 261
Qy 23 VDKNT 28

RESULT 10
ENTRY
#molecule_type DNA
#type complete
TITLE
beta-glucosidase homolog - Lyme disease spirochete
ORGANISM
#formal_name Borrelia burgdorferi #common_name Lyme disease
        spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
05-Jun-1998
ACCESSIONS
REFERENCE
#authors
Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Utterback, T.; Watthey, L.; McDonald, L.; Artiach, P.;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
Nature (1997) 390:580-586
#journal
#title Genomic sequence of a Lyme disease spirochaete, Borrelia
        burgdorferi.
#cross-references MUID:98065943
#accession
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-555 #label KLE
#cross-references GB:AE001163; GB:AE000783; NID:g2688541; PID:g2688545;
        TIGR:BB0620
#experimental_source strain B31
SUMMARY
#length 555 #molecular-weight 63964 #checksum 9643
Query Match 20.7%; Score 6; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 199 GQKGG 204
Qy 16 GQKGG 21

RESULT 11
ENTRY
#molecule_type DNA
#type complete
TITLE
hypothetical protein ag_2054 - Aquifex aeolicus
ORGANISM
#formal_name Aquifex aeolicus
DATE 08-May-1998 #sequence_revision 08-May-1998 #text_change

```

```

08-May-1998
ACCESSIONS B70476
REFERENCE A70300
#authors Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;
          Lennox, A.L.; Graham, D.E.; Overbeek, R.; Sneed, M.A.;
          Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short,
          J.M.; Olson, G.J.; Swanson, R.V.
#journal Nature (1998) 392:353-358
#title The complete genome of the hyperthermophilic bacterium
        Aquifex aeolicus.
#cross-references MUID:98196666
#accession B70476
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
##residues 1-1116 ##label AQF
##cross-references GB:AE000770; NID:g2984274; PID:g2984280; GB:AE000657
##experimental_source strain VF5
GENETICS
#gene aq_2054
SUMMARY #length 1116 #molecular-weight 127794 #checksum 5967
Query Match 20.7%; Score 6; DB 2; Length 1116;
Best Local Similarity 100.0%; Pred. No. 3.54e+00; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Db 289 VVDKNT 294
QY 23 VVDKNT 28
|||||

RESULT 12
ENTRY S27224 #type complete
TITLE N-methyl-D-aspartate receptor epsilon-4 chain - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change
17-Mar-1999
ACCESSIONS S27224
REFERENCE S27224
#authors Ikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.;
          Watanabe, M.; Inoue, Y.; Mishina, M.
#journal FEBS Lett. (1992) 313:34-38
#title Cloning and expression of the epsilon-4 subunit of the NMDA
        receptor channel.
#cross-references MUID:93050214
#accession S27224
#status preliminary
#molecule_type mRNA
##residues 1-1323 ##label IKE
##cross-references EMBL:D12822
CLASSIFICATION #superfamily N-methyl-D-aspartate receptor 2D; glutamate
        receptor homology
KEYWORDS transmembrane protein
FEATURE
451-879 #domain glutamate receptor homology #label GRH
SUMMARY #length 1323 #molecular-weight 143301 #checksum 3104
Query Match 20.7%; Score 6; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 991 TTQPPQ 996
QY 3 TTQPPQ 8
|||||

RESULT 13
ENTRY I78557 #type complete
TITLE N-methyl-D-aspartate receptor chain NMDAR2D-2 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
05-Dec-1997
ACCESSIONS I78557; I58158; D45219

```

```

I58158
REFERENCE Monyer, H.; Burnashev, N.; Laurie, D.J.; Sakmann, B.;
          Seeburg, P.H.
#journal Neuron (1994) 12:529-540
#title Developmental and regional expression in the rat brain and
        functional properties of four NMDA receptors.
#cross-references MUID:94206533
#accession I78557
#status preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
##residues 1-1323 ##label RES
##cross-references GB:L31612; NID:g469068; PID:g469069
#accession I58158
#status preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
##residues 1-66,'V',68-1323 ##label RE2
##cross-references GB:L31611; NID:g469066; PID:g469067
REFERENCE A45219
#authors Ishii, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.;
          Kadotani, H.; Yokoi, M.; Akazawa, C.; Shigemoto, R.;
          Mizuno, N.; Masu, M.; Nakanishi, S.
#journal J. Biol. Chem. (1993) 268:2836-2843
#title Molecular characterization of the family of the
        N-methyl-D-aspartate receptor subunits.
#cross-references MUID:93155102
#accession D45219
#status preliminary; not compared with conceptual translation
#molecule_type nucleic acid
##residues 1265-1323 ##label ISH
##experimental_source brain
##note sequence extracted from NCBI backbone (NCBIP:124265)
        receptor homology
CLASSIFICATION #superfamily N-methyl-D-aspartate receptor 2D; glutamate
        receptor homology
SUMMARY #length 1323 #molecular-weight 142956 #checksum 3610
Query Match 20.7%; Score 6; DB 2; Length 1323;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 991 TTQPPQ 996
QY 3 TTQPPQ 8
|||||

RESULT 14
ENTRY C45219 #type complete
TITLE N-methyl-D-aspartate receptor chain NMDAR2D-1 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change
12-May-1995
ACCESSIONS C45219
REFERENCE C45219
#authors Ishii, T.; Moriyoshi, K.; Sugihara, H.; Sakurada, K.;
          Kadotani, H.; Yokoi, M.; Akazawa, C.; Shigemoto, R.;
          Mizuno, N.; Masu, M.; Nakanishi, S.
#journal J. Biol. Chem. (1993) 268:2836-2843
#title Molecular characterization of the family of the
        N-methyl-D-aspartate receptor subunits.
#cross-references MUID:93155102
#accession C45219
#status preliminary; not compared with conceptual translation
#molecule_type nucleic acid
##residues 1-1356 ##label ISH
##experimental_source brain
##note sequence extracted from NCBI backbone (NCBIP:124264)
        receptor homology
CLASSIFICATION #superfamily N-methyl-D-aspartate receptor 2D; glutamate
        receptor homology
FEATURE
451-879 #domain glutamate receptor homology #label GRH
SUMMARY #length 1356 #molecular-weight 145973 #checksum 5329
Query Match 20.7%; Score 6; DB 2; Length 1356;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;

```

```

Matches      6:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db   991  TTQPPQ 996
      |||||
Qy    3  TTQPPQ 8

RESULT      15
ENTRY      B47447      #type complete
TITLE      calcium channel protein alpha-1 chain (variant doe-4) -
            electric ray (Discopyge ommata)
ORGANISM   #formal_name Discopyge ommata
DATE       21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change
            13-Sep-1998
ACCESSIONS B47447
REFERENCE  A47447
            Horne, W.A.; Ellinor, P.T.; Inman, I.; Zhou, M.; Tsien, R.W.;
            Schwarz, T.L.
            Proc. Natl. Acad. Sci. U.S.A. (1993) 90:3787-3791
            Molecular diversity of Ca(2+) channel alpha 1 subunits from
            the marine ray Discopyge ommata.
            #cross-references MUID:93248175
            #accession B47447
            ##status preliminary; not compared with conceptual translation
            ##molecule_type mRNA
            ##residues 1-2326 ##label HOR
            ##note sequence extracted from NCBI backbone (NCBIP:130673)
CLASSIFICATION #superfamily voltage-dependent calcium channel protein
                alpha-1 chain
SUMMARY        #length 2326 #molecular-weight 264515 #checksum 2898

Query Match      20.7%; Score 6; DB 2; Length 2326;
Best Local Similarity 100.0%; Pred. No. 3.54e+00;
Matches      6:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db   2284 TTQPPQ 2289
      |||||
Qy    3  TTQPPQ 8

```

Search completed: Sat Aug 28 14:54:03 1999  
 Job time : 43 secs.

(TM)

Result No.	Query		Length	DB	ID	Description	Pred. No.
	Match	Score					
1	7	24.1	999	1	DSG3_HUMAN	DESMOGLEIN 3 PRECURSOR	2.08e-02
2	6	20.7	124	1	VATF_MANSE	VACUOLAR ATP SYNTHASE	1.67e+00
3	6	20.7	124	1	VATF_DROME	VACUOLAR ATP SYNTHASE	1.67e+00
4	6	20.7	127	1	VATF_ANOGA	VACUOLAR ATP SYNTHASE	1.67e+00
5	6	20.7	132	1	SUCA_HUMAN	SUCCINYL-COA LIGASE [G	1.67e+00
6	6	20.7	154	1	AR16_YEAST	ARP2/3 COMPLEX 16 KD S	1.67e+00
7	6	20.7	205	1	FLRE_BOVIN	FLAVIN REDUCTASE (EC 1	1.67e+00
8	6	20.7	264	1	LPXA_RICRI	ACYL-[ACYL-CARRIER-PRO	1.67e+00
9	6	20.7	508	1	FTS2_CAUCR	CELL DIVISION PROTEIN	1.67e+00
10	6	20.7	596	1	SDP_EIMBO	SPOROZOITE DEVELOPMENT	1.67e+00
11	6	20.7	630	1	FIG1_MOUSE	FIG-1 PROTEIN PRECURSOR	1.67e+00
12	6	20.7	782	1	CAD5_PIG	VASCULAR ENDOTHELIAL-C	1.67e+00
13	6	20.7	1323	1	NME4_MOUSE	GLUTAMATE [NMDA] RECEP	1.67e+00
14	5	17.2	76	1	FIXU_RHISN	FIXU PROTEIN HOMOLOG.	8.42e+01
15	5	17.2	84	1	SCX2_CENNO	TOXIN 2 PRECURSOR [TOX	8.42e+01
16	5	17.2	110	1	RL24_THETH	50S RIBOSOMAL PROTEIN	8.42e+01
17	5	17.2	118	1	YTON_ENTAE	HYPOTHETICAL PROTEIN I	8.42e+01
18	5	17.2	121	1	FLIT_ECOLI	FLAGELLAR PROTEIN FLIT	8.42e+01
19	5	17.2	123	1	YTON_KLEPN	HYPOTHETICAL PROTEIN I	8.42e+01
20	5	17.2	129	1	CYC3_DESVH	CYTOCHROME C3 PRECURSOR	8.42e+01
21	5	17.2	138	1	Y922_METJA	HYPOTHETICAL PROTEIN M	8.42e+01
22	5	17.2	143	1	TS22_MOUSE	PUTATIVE REGULATORY PR	8.42e+01
23	5	17.2	144	1	TS22_HUMAN	PUTATIVE REGULATORY PR	8.42e+01

97	5	17.2	382	1	FTSZ_BACSU	CELL DIVISION PROTEIN	8.42e+01	170	5	17.2	707	1	TRFL_MOUSE	LACTOTRANSFERRIN PRECU	8.42e+01
98	5	17.2	385	1	CEB3_CAEEL	G2/MITOTIC-SPECIFIC CY	8.42e+01	171	5	17.2	710	1	PET1_RAT	OLIGOPEPTIDE TRANSPORT	8.42e+01
99	5	17.2	386	1	YMK7_YEAST	HYPOTHETICAL 45.2 KD P	8.42e+01	172	5	17.2	711	1	TRFL_HUMAN	LACTOTRANSFERRIN PRECU	8.42e+01
100	5	17.2	393	1	NCAP_PVM	NUCLEOCAPSID PROTEIN	8.42e+01	173	5	17.2	720	1	ABI3_ARATH	ABSCISIC ACID-INSENSIT	8.42e+01
101	5	17.2	393	1	GUN1_USTWA	ENDOGLUCANASE 1 PRECUR	8.42e+01	174	5	17.2	725	1	IF3X_SCHPO	PROBABLE EUKARYOTIC TR	8.42e+01
102	5	17.2	393	1	PG12_ERWCA	ENDO-POLYGLACTURONASE	8.42e+01	175	5	17.2	738	1	NUAM_SOLITU	NADH-UBIQUINONE OXIDOR	8.42e+01
103	5	17.2	403	1	CAG1_MOUSE	CDP-N-ACETYLNEURAMINAT	8.42e+01	176	5	17.2	738	1	PAP_BOVIN	POLY(A) POLYMERASE (EC	8.42e+01
104	5	17.2	414	1	LMP1_CHICK	LYSOSOME-ASSOCIATED ME	8.42e+01	177	5	17.2	743	1	EGAL_THEET	BETA-GALACTOSIDASE (EC	8.42e+01
105	5	17.2	420	1	Y181_MYCGE	HYPOTHETICAL PROTEIN M	8.42e+01	178	5	17.2	753	1	PPSA_ARCFU	PROBABLE PHOSPHOENOLPY	8.42e+01
106	5	17.2	423	1	GATM_HUMAN	GLYCINE AMIDNOTRANSFE	8.42e+01	179	5	17.2	763	1	IF3X_YEAST	EUKARYOTIC TRANSLATION	8.42e+01
107	5	17.2	436	1	EP45_XENLA	ESTROGEN-REGULATED PRO	8.42e+01	180	5	17.2	783	1	CAD5_MOUSE	VASCULAR ENDOTHELIAL-C	8.42e+01
108	5	17.2	441	1	TBB7_CAEEL	TUBULIN BETA CHAIN	8.42e+01	181	5	17.2	784	1	YO63_CAEEL	HYPOTHETICAL 92.1 KD P	8.42e+01
109	5	17.2	444	1	NUDF_EMENT	NUCLEAR MIGRATION PROT	8.42e+01	182	5	17.2	788	1	UGS3_SOLITU	GLYCOPEN [STARCH] STNT	8.42e+01
110	5	17.2	444	1	GLPT_BACSU	GLYCEROL-3-PHOSPHATE T	8.42e+01	183	5	17.2	807	1	MCM3_XENLA	DNA REPLICATION LICENS	8.42e+01
111	5	17.2	445	1	Y148_MYCPN	HYPOTHETICAL PROTEIN M	8.42e+01	184	5	17.2	819	1	UL47_VZVD	ALPHA TRANS-INDUCING F	8.42e+01
112	5	17.2	445	1	HGD_HUMAN	HOMOGENTISATE 1,2-DIOX	8.42e+01	185	5	17.2	821	1	CTR1_ARATH	SERINE/THREONINE-PROTE	8.42e+01
113	5	17.2	448	1	EY2B_HUMAN	EV12B PROTEIN PRECURSO	8.42e+01	186	5	17.2	827	1	GYRA_HELPY	DNA GYRASE SUBUNIT A (	8.42e+01
114	5	17.2	450	1	VG40_HSVSA	HYPOTHETICAL GENE 40 P	8.42e+01	187	5	17.2	845	1	MAT3_RAT	MATRIN 3	8.42e+01
115	5	17.2	459	1	ME31_DROME	PUTATIVE ATP-DEPENDENT	8.42e+01	188	5	17.2	858	1	CNRA_BOVIN	ROD CGMP-SPECIFIC 3',5	8.42e+01
116	5	17.2	463	1	NARK_ECOLI	NITRATE EXTRUSION PROT	8.42e+01	189	5	17.2	858	1	CNRA_MOUSE	ROD CGMP-SPECIFIC 3',5	8.42e+01
117	5	17.2	467	1	VG3H_YEAST	HYPOTHETICAL 54.5 KD P	8.42e+01	190	5	17.2	859	1	CNRA_HUMAN	ROD CGMP-SPECIFIC 3',5	8.42e+01
118	5	17.2	468	1	LSU2_HAEIN	3-ISOPROPYLMALATE DEHY	8.42e+01	191	5	17.2	860	1	CNRA_CANFA	ROD CGMP-SPECIFIC 3',5	8.42e+01
119	5	17.2	469	1	Y889_METJA	HYPOTHETICAL PROTEIN M	8.42e+01	192	5	17.2	868	1	VGLB_VZVD	GLYCOPROTEIN B PRECURS	8.42e+01
120	5	17.2	474	1	VP61_NVPOP	61 KD PROTEIN HOMOLOG	8.42e+01	193	5	17.2	878	1	YD97_SCHPO	HYPOTHETICAL 98.4 KD P	8.42e+01
121	5	17.2	475	1	Y148_YEAST	HYPOTHETICAL 52.6 KD P	8.42e+01	194	5	17.2	886	1	VGP3_EBV8	ENVELOPE GLYCOPROTEIN	8.42e+01
122	5	17.2	477	1	Y142_NPVAC	HYPOTHETICAL 55.4 KD P	8.42e+01	195	5	17.2	907	1	VGP3_EBV	26S PROTEASOME REGULAT	8.42e+01
123	5	17.2	477	1	P3_HUMAN	P3 PROTEIN	8.42e+01	196	5	17.2	909	1	PSD2_HUMAN	PROBABLE TRANSFERRIN B	8.42e+01
124	5	17.2	478	1	RPB1_EUPOC	DNA-DIRECTED RNA POLYM	8.42e+01	197	5	17.2	912	1	TBP1_HAEIN	PROBABLE TRANSFERRIN B	8.42e+01
125	5	17.2	478	1	NAOX_MYCGE	PROBABLE NADH OXIDASE	8.42e+01	198	5	17.2	913	1	VGLB_PRIVF	GLYCOPROTEIN GII PRECU	8.42e+01
126	5	17.2	488	1	RBL_ANTSP	RIBULOSE BISPHOSPHATE	8.42e+01	199	5	17.2	917	1	SYL_CAMEJ	ISOLEUCYL-TRNA SYNTHET	8.42e+01
127	5	17.2	495	1	NU4M_MARPO	NADH-UBIQUINONE OXIDOR	8.42e+01	200	5	17.2	923	1	YAU4_SCHPO	HYPOTHETICAL 104.5 KD	8.42e+01
128	5	17.2	495	1	NU4M_WHEAT	NADH-UBIQUINONE OXIDOR	8.42e+01	201	5	17.2	928	1	VGLB_HSVBP	GLYCOPROTEIN I PRECURS	8.42e+01
129	5	17.2	496	1	YKAB_CAEEL	HYPOTHETICAL 55.0 KD P	8.42e+01	202	5	17.2	932	1	VGLB_HSVBP	GLYCOPROTEIN I PRECURS	8.42e+01
130	5	17.2	501	1	MYSU_RABIT	MYOSIN HEAVY CHAIN, EM	8.42e+01	203	5	17.2	932	1	PMS1_HUMAN	PMS1 PROTEIN HOMOLOG 1	8.42e+01
131	5	17.2	504	1	COA3_NAV2	PROBABLE COAT PROTEIN	8.42e+01	204	5	17.2	943	1	DIP2_YEAST	DOM34 INTERACTING PROT	8.42e+01
132	5	17.2	504	1	CPK1_ONCMY	CYTCHROME P450 2K1 (E	8.42e+01	205	5	17.2	948	1	UVRA_MYCPN	EXCINUCLEASE ABC SUBUN	8.42e+01
133	5	17.2	510	1	YJY5_YEAST	HYPOTHETICAL 58.1 KD P	8.42e+01	206	5	17.2	952	1	UVRA_MYCGE	EXCINUCLEASE ABC SUBUN	8.42e+01
134	5	17.2	510	1	GATA_USTWA	4-AMINOBUSTRATE AMINOT	8.42e+01	207	5	17.2	980	1	POLG_LIV	GENOME POLYPROTEIN (CO	8.42e+01
135	5	17.2	515	1	NCAP_P13H4	NUCLEOCAPSID PROTEIN	8.42e+01	208	5	17.2	997	1	YFM3_YEAST	TRANSLATION INITIATION	8.42e+01
136	5	17.2	520	1	CYCB_ALCEU	COBALT-ZINC-CADMIUM RE	8.42e+01	209	5	17.2	1002	1	IF2P_YEAST	TRANSLATION INITIATION	8.42e+01
137	5	17.2	520	1	CYCB_ALCEP	CATION EFFLUX SYSTEM P	8.42e+01	210	5	17.2	1023	1	EGAL_ECOLI	BETA-GALACTOSIDASE (EC	8.42e+01
138	5	17.2	522	1	STP1_ARATH	GLUCOSE TRANSPORTER (S	8.42e+01	211	5	17.2	1025	1	YEGO_ECOLI	HYPOTHETICAL 111.0 KD	8.42e+01
139	5	17.2	527	1	YRIP_ECOLI	HYPOTHETICAL 59.7 KD P	8.42e+01	212	5	17.2	1028	1	EGAL_ENTCL	BETA-GALACTOSIDASE (EC	8.42e+01
140	5	17.2	535	1	YB30_METJA	HYPOTHETICAL PROTEIN M	8.42e+01	213	5	17.2	1033	1	TIR1_ECOLI	TYPE I RESTRICTION ENZ	8.42e+01
141	5	17.2	536	1	GAG_MLVCB	GAG POLYPROTEIN (CONTA	8.42e+01	214	5	17.2	1034	1	GCSB_FLAPR	GLYCINE DEHYDROGENASE	8.42e+01
142	5	17.2	538	1	GAG_MLVMO	GAG POLYPROTEIN (CONTA	8.42e+01	215	5	17.2	1034	1	GCSF_FLATR	GLYCINE DEHYDROGENASE	8.42e+01
143	5	17.2	547	1	SNGL_YEAST	NITROSOGUANIDINE RESIS	8.42e+01	216	5	17.2	1034	1	GCSF_FLAPR	GLYCINE DEHYDROGENASE	8.42e+01
144	5	17.2	575	1	PRTD_ERWCH	PROTEASES SECRETION AT	8.42e+01	217	5	17.2	1037	1	GCSA_FLAPR	GLYCINE DEHYDROGENASE	8.42e+01
145	5	17.2	583	1	C166_HUMAN	CD166 ANTIGEN PRECURSO	8.42e+01	218	5	17.2	1040	1	YEGN_ECOLI	HYPOTHETICAL 112.1 KD	8.42e+01
146	5	17.2	587	1	COAT_PAVL3	COAT PROTEIN VP1 (CONT	8.42e+01	219	5	17.2	1090	1	PULA_KLEPN	PULLULANASE PRECURSOR	8.42e+01
147	5	17.2	587	1	ASO_CUCSA	L-ASCORBATE OXIDASE PR	8.42e+01	220	5	17.2	1120	1	YB85_YEAST	HYPOTHETICAL 124.0 KD	8.42e+01
148	5	17.2	588	1	C166_CHICK	CD166 ANTIGEN PRECURSO	8.42e+01	221	5	17.2	1120	1	MYSB_ACACA	MYOSIN HEAVY CHAIN IB	8.42e+01
149	5	17.2	589	1	MAZS_YEAST	PROBABLE ALPHA-GLUCOSI	8.42e+01	222	5	17.2	1163	1	ITAX_HUMAN	LEUKOCYTE ADHESION GLY	8.42e+01
150	5	17.2	603	1	ALU4_HUMAN	!!!! ALU SUBFAMILY SB2	8.42e+01	223	5	17.2	1184	1	FBL2_HUMAN	FIBULIN-2 PRECURSOR	8.42e+01
151	5	17.2	610	1	PR2A_ARATH	PEPTIDE TRANSPORTER PT	8.42e+01	224	5	17.2	1216	1	AEGP_RAT	APICAL ENDOSOMAL GLYCO	8.42e+01
152	5	17.2	612	1	DHGL_DROPS	GLUCOSE DEHYDROGENASE	8.42e+01	225	5	17.2	1221	1	FBL2_MOUSE	FIBULIN-2 PRECURSOR	8.42e+01
153	5	17.2	636	1	FTHS_SPIOI	FORMATE--TETRAHYDROFOL	8.42e+01	226	5	17.2	1231	1	CFAH_HUMAN	COMPLEMENT FACTOR H PR	8.42e+01
154	5	17.2	637	1	RES1_SCHPO	CELL DIVISION CYCLE RE	8.42e+01	227	5	17.2	1244	1	SLA1_YEAST	CYTOSKELETON ASSEMBLY	8.42e+01
155	5	17.2	639	1	V70K_PLRVL	69.7 KD PROTEIN (ORF 2	8.42e+01	228	5	17.2	1254	1	MDR3_CAEEL	MULTIDRUG RESISTANCE P	8.42e+01
156	5	17.2	639	1	V70K_PLRVL	69.7 KD PROTEIN (ORF 2	8.42e+01	229	5	17.2	1255	1	MUC1_HUMAN	MUCIN 1 PRECURSOR (POL	8.42e+01
157	5	17.2	646	1	G811_CANAL	GL/S-SPECIFIC CYCLIN C	8.42e+01	230	5	17.2	1279	1	MDR3_HUMAN	MULTIDRUG RESISTANCE P	8.42e+01
158	5	17.2	646	1	SXBL_SCHPO	SHK1 KINASE-BINDING PR	8.42e+01	231	5	17.2	1286	1	TRBP_BOVIN	INTERPHORECEPTOR RET	8.42e+01
159	5	17.2	652	1	PAP_HUMAN	POLY(A) POLYMERASE (EC	8.42e+01	232	5	17.2	1380	1	ZMS1_YEAST	ZINC FINGER PROTEIN ZM	8.42e+01
160	5	17.2	656	1	RAF2_HUMAN	RAB PROTEINS GERANYLGE	8.42e+01	233	5	17.2	1407	1	T2D2_YEAST	TRANSCRIPTION INITIATI	8.42e+01
161	5	17.2	656	1	HFG_MICLE	HEAT SHOCK PROTEIN HTP	8.42e+01	234	5	17.2	1520	1	GLTB_BACSU	GLUTAMATE SYNTHASE [NA	8.42e+01
162	5	17.2	656	1	TOP3_YEAST	DNA TOPOISOMERASE III	8.42e+01	235	5	17.2	1616	1	SLAP_BACCI	S-LAYER RELATED PROTEI	8.42e+01
163	5	17.2	660	1	Y390_MYCGE	HYPOTHETICAL ATP-BINDI	8.42e+01	236	5	17.2	1722	1	YP93_CAEEL	HYPOTHETICAL HELICASE	8.42e+01
164	5	17.2	660	1	Y390_MYCPN	HYPOTHETICAL ATP-BINDI	8.42e+01	237	5	17.2	1776	1	POLR_OYMW	RNA REPLICASE POLYPROT	8.42e+01
165	5	17.2	662	1	GARP_HUMAN	GARP PROTEIN PRECURSOR	8.42e+01	238	5	17.2	1839	1	CYAA_SACKL	ADENYLATE CYCLASE (EC	8.42e+01
166	5	17.2	666	1	YEA7_YEAST	HYPOTHETICAL 72.5 KD P	8.42e+01	239	5	17.2	1867	1	MOT1_YEAST	PROBABLE HELICASE MOT1	8.42e+01
167	5	17.2	673	1	YU18_CAEEL	PUTATIVE ACETYLCHOLINE	8.42e+01	240	5	17.2	1955	1	AGRI_CHICK	AGRI PRECURSOR	8.42e+01
168	5	17.2	677	1	SACC_BACSU	LEVANASE PRECURSOR (EC	8.42e+01	241	5	17.2	1976	1	MYSO_HUMAN	MYOSIN HEAVY CHAIN, NO	8.42e+01
169	5	17.2	680	1	SYG_BOMMO	GLCYL-TRNA SYNTHETASE	8.42e+01	242	5	17.2	2280	1	YCF2_TOBAC	HYPOTHETICAL 267 KD PR	8.42e+01

243	5	17.2	2291	1	RRPB_BEV	RNA-DIRECTED RNA POLYM	8.42e+01	315	4	13.8	158	1	YCBM_BACSU	HYPOTHETICAL 18.0 KD P	2.22e+03
244	5	17.2	2376	1	YIM5_YEAST	HYPOTHETICAL 269.9 KD	8.42e+01	317	4	13.8	159	1	Y215_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
245	5	17.2	2505	1	FAS_RAT	FATTY ACID SYNTHASE (E	8.42e+01	318	4	13.8	160	1	YHVI_LACHE	HYPOTHETICAL PROTEIN I	2.22e+03
246	5	17.2	2670	1	YAO5_SCHPO	PUTATIVE TRANSLATIONAL	8.42e+01	319	4	13.8	162	1	YPU4_RHOCA	HYPOTHETICAL 17.4 KD P	2.22e+03
247	5	17.2	3256	1	KI67_HUMAN	ANTIGEN KI-67	8.42e+01	320	4	13.8	163	1	YA9F_SCHPO	HYPOTHETICAL 18.1 KD P	2.22e+03
248	5	17.2	3433	1	POLG_KUNJM	GENOME POLYPROTEIN [CO	8.42e+01	321	4	13.8	165	1	Y243_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
249	5	17.2	4540	1	DYHC_PARTE	DYNEIN HEAVY CHAIN, CY	8.42e+01	322	4	13.8	165	1	Y4UD_RHISN	HYPOTHETICAL 18.7 KD P	2.22e+03
250	5	17.2	4568	1	DYHC_CABEL	DYNEIN HEAVY CHAIN, CY	8.42e+01	323	4	13.8	165	1	YAJ5_SCHPO	ADAPTIN COMPLEX SMALL	2.22e+03
251	5	17.2	4639	1	DYHC_DROME	DYNEIN HEAVY CHAIN, CY	8.42e+01	324	4	13.8	169	1	YMO3_MARPO	HYPOTHETICAL 19.3 KD P	2.22e+03
252	5	17.2	4644	1	DYHC_RAT	DYNEIN HEAVY CHAIN, CY	8.42e+01	325	4	13.8	171	1	YLY1_STAUA	HYPOTHETICAL 18.6 KD P	2.22e+03
253	5	17.2	4725	1	DYHC_DICDI	DYNEIN HEAVY CHAIN, CY	8.42e+01	326	4	13.8	173	1	Y028_NPVAC	HYPOTHETICAL 20.4 KD P	2.22e+03
254	4	13.8	50	1	YORJ_TTV1	HYPOTHETICAL 6.1 KD PR	2.22e+03	327	4	13.8	174	1	YC21_PORPU	HYPOTHETICAL 20.6 KD P	2.22e+03
255	4	13.8	54	1	VRPI_BPT3	BACTERIAL RNA POLYMER	2.22e+03	328	4	13.8	174	1	YI19_HUMAN	HYPOTHETICAL PROTEIN M	2.22e+03
256	4	13.8	55	1	YPI8_CLOPE	HYPOTHETICAL 6.1 KD PR	2.22e+03	329	4	13.8	174	1	YA68_METTH	HYPOTHETICAL PROTEIN M	2.22e+03
257	4	13.8	55	1	Y085_NPVOP	HYPOTHETICAL 6.1 KD PR	2.22e+03	330	4	13.8	176	1	YGDP_ECOLI	HYPOTHETICAL 20.8 KD P	2.22e+03
258	4	13.8	60	1	Y04F_BPT4	HYPOTHETICAL 7.0 KD PR	2.22e+03	331	4	13.8	176	1	Y094_MYCTU	HYPOTHETICAL 19.9 KD P	2.22e+03
259	4	13.8	64	1	YMF8_ECOLI	HYPOTHETICAL 6.4 KD PR	2.22e+03	332	4	13.8	177	1	YAL6_SCHPO	HYPOTHETICAL 20.1 KD P	2.22e+03
260	4	13.8	64	1	YPT1_STRCO	HYPOTHETICAL 6.4 KD PR	2.22e+03	333	4	13.8	178	1	Y057_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
261	4	13.8	65	1	YVEE_VACCC	HYPOTHETICAL 7.6 KD PR	2.22e+03	334	4	13.8	179	1	YAF4_CAEEL	HYPOTHETICAL 20.4 KD P	2.22e+03
262	4	13.8	68	1	Y055_NPVOP	HYPOTHETICAL 8.0 KD PR	2.22e+03	335	4	13.8	180	1	YAF5_YEAST	HYPOTHETICAL 19.9 KD P	2.22e+03
263	4	13.8	72	1	VOGR_BBP2	LATE CONTROL OGR PROTE	2.22e+03	336	4	13.8	180	1	YFXK_BRAJA	HYPOTHETICAL PROTEIN I	2.22e+03
264	4	13.8	77	1	Y660_METJA	HYPOTHETICAL 8.7 KD PR	2.22e+03	337	4	13.8	181	1	YK69_CAEEL	FIMBRIAL-LIKE PROTEIN	2.22e+03
265	4	13.8	80	1	YOT0_CABEL	HYPOTHETICAL 8.7 KD PR	2.22e+03	338	4	13.8	183	1	YGIL_ECOLI	HYPOTHETICAL 20.9 KD P	2.22e+03
266	4	13.8	81	1	Y154_NPVAC	HYPOTHETICAL 9.4 KD PR	2.22e+03	339	4	13.8	184	1	YPS2_SYN2	HYPOTHETICAL 20.5 KD P	2.22e+03
267	4	13.8	84	1	Y0AF_ECOLI	HYPOTHETICAL 8.9 KD PR	2.22e+03	340	4	13.8	185	1	YAIS_ECOLI	HYPOTHETICAL 21.5 KD P	2.22e+03
268	4	13.8	89	1	Y2FA_ECOLI	VERY HYPOTHETICAL 10.6	2.22e+03	341	4	13.8	186	1	YP89_CAEEL	HYPOTHETICAL 20.3 KD P	2.22e+03
269	4	13.8	94	1	YIF8_YEAST	HYPOTHETICAL 10.3 KD P	2.22e+03	342	4	13.8	186	1	YHBO_ECOLI	HYPOTHETICAL PROTEIN M	2.22e+03
270	4	13.8	95	1	Y121_CLOPE	HYPOTHETICAL 10.7 KD P	2.22e+03	343	4	13.8	187	1	YG61_METJA	HYPOTHETICAL 20.8 KD P	2.22e+03
271	4	13.8	97	1	YBGE_ECOLI	10.9 KD PROTEIN IN CYD	2.22e+03	344	4	13.8	187	1	Y034_BHP1	HYPOTHETICAL 20.8 KD P	2.22e+03
272	4	13.8	99	1	YIIS_ECOLI	HYPOTHETICAL 10.8 KD P	2.22e+03	345	4	13.8	189	1	Y031_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
273	4	13.8	101	1	YGE2_YEAST	HYPOTHETICAL 12.1 KD P	2.22e+03	346	4	13.8	193	1	YAK7_YEAST	VERY HYPOTHETICAL 21.1	2.22e+03
274	4	13.8	103	1	YPC6_CLAUP	HYPOTHETICAL 11.4 KD P	2.22e+03	347	4	13.8	194	1	Y248_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
275	4	13.8	104	1	Y237_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	348	4	13.8	196	1	YGDP_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
276	4	13.8	104	1	YK12_YEAST	HYPOTHETICAL 12.4 KD P	2.22e+03	349	4	13.8	197	1	WRBA_ECOLI	TRP REPRESSOR BINDING	2.22e+03
277	4	13.8	108	1	Y04B_BPT4	HYPOTHETICAL 12.7 KD P	2.22e+03	350	4	13.8	197	1	Y882_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
278	4	13.8	108	1	YN11_CAEEL	HYPOTHETICAL 12.1 KD P	2.22e+03	351	4	13.8	199	1	YE74_HAEIN	HYPOTHETICAL ABC TRANS	2.22e+03
279	4	13.8	109	1	YXIH_BACSU	HYPOTHETICAL 12.4 KD P	2.22e+03	352	4	13.8	199	1	Y143_NBPV	HYPOTHETICAL PROTEIN B	2.22e+03
280	4	13.8	110	1	Y225_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	353	4	13.8	201	1	Y146_NPVAC	HYPOTHETICAL 22.9 KD P	2.22e+03
281	4	13.8	113	1	Y08B_MYCTU	HYPOTHETICAL 11.6 KD P	2.22e+03	354	4	13.8	201	1	Y188_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
282	4	13.8	114	1	YCX7_PORPU	HYPOTHETICAL 12.4 KD P	2.22e+03	355	4	13.8	202	1	Y009_PYRHO	HYPOTHETICAL PROTEIN P	2.22e+03
283	4	13.8	115	1	Y4HO_RHISN	HYPOTHETICAL 12.8 KD P	2.22e+03	356	4	13.8	202	1	Y099_CAEEL	HYPOTHETICAL 22.7 KD P	2.22e+03
284	4	13.8	115	1	YTR3_SPAU	HYPOTHETICAL 12.7 KD P	2.22e+03	357	4	13.8	202	1	Y063_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
285	4	13.8	116	1	Y987_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	358	4	13.8	203	1	YBM4_YEAST	HYPOTHETICAL 22.6 KD P	2.22e+03
286	4	13.8	117	1	YBA4_ECOLI	HYPOTHETICAL 13.3 KD P	2.22e+03	359	4	13.8	203	1	YPUJ_BACSU	HYPOTHETICAL 23.7 KD P	2.22e+03
287	4	13.8	118	1	YDF4_SCHPO	VERY HYPOTHETICAL 14.1	2.22e+03	360	4	13.8	203	1	YJ44_YEAST	HYPOTHETICAL 23.7 KD P	2.22e+03
288	4	13.8	119	1	YB82_YEAST	HYPOTHETICAL 13.9 KD P	2.22e+03	361	4	13.8	205	1	YAB4_YEAST	HYPOTHETICAL 22.4 KD P	2.22e+03
289	4	13.8	120	1	YWLA_BACSU	HYPOTHETICAL 13.9 KD P	2.22e+03	362	4	13.8	205	1	YDDH_ECOLI	HYPOTHETICAL 23.7 KD P	2.22e+03
290	4	13.8	121	1	YGIN_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	363	4	13.8	205	1	YB23_METJA	HYPOTHETICAL 22.8 KD P	2.22e+03
291	4	13.8	124	1	YQPS_CABEL	HYPOTHETICAL 13.8 KD P	2.22e+03	364	4	13.8	206	1	YFGN_ECOLI	HYPOTHETICAL 22.2 KD P	2.22e+03
292	4	13.8	124	1	YF80_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	365	4	13.8	206	1	Y930_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
293	4	13.8	124	1	YEEV_ECOLI	HYPOTHETICAL 13.9 KD P	2.22e+03	366	4	13.8	206	1	Y312_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
294	4	13.8	127	1	Y389_MYCGE	HYPOTHETICAL 13.9 KD P	2.22e+03	367	4	13.8	206	1	Y615_SYNY3	HYPOTHETICAL 22.4 KD P	2.22e+03
295	4	13.8	128	1	YPRR_ECOLI	HYPOTHETICAL PROTEIN R	2.22e+03	368	4	13.8	206	1	YMA8_BACSU	HYPOTHETICAL 23.4 KD P	2.22e+03
296	4	13.8	128	1	YJ22_YEAST	HYPOTHETICAL 14.5 KD P	2.22e+03	369	4	13.8	208	1	Y086_CAEEL	HYPOTHETICAL 22.9 KD P	2.22e+03
297	4	13.8	130	1	YGIN_ECOLI	11.9 KD PROTEIN IN PAR	2.22e+03	370	4	13.8	209	1	Y379_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
298	4	13.8	132	1	Y334_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	371	4	13.8	209	1	YAC2_SCHPO	HYPOTHETICAL 23.9 KD P	2.22e+03
299	4	13.8	140	1	Y337_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03	372	4	13.8	209	1	Y222_PORPU	HYPOTHETICAL 23.7 KD P	2.22e+03
300	4	13.8	144	1	WAP3_PIG	WAP-3 PROTEIN PRECURSO	2.22e+03	373	4	13.8	211	1	YH11_LACLA	HYPOTHETICAL PROTEIN I	2.22e+03
301	4	13.8	146	1	YDPI_SULAC	HYPOTHETICAL 16.4 KD P	2.22e+03	374	4	13.8	211	1	YH12_SOLTU	WOUND-INDUCED PROTEIN	2.22e+03
302	4	13.8	146	1	YN68_YEAST	HYPOTHETICAL 16.5 KD P	2.22e+03	375	4	13.8	214	1	Y540_AQUAE	HYPOTHETICAL PROTEIN A	2.22e+03
303	4	13.8	146	1	Y146_LAMBO	HYPOTHETICAL NIN REGIO	2.22e+03	376	4	13.8	215	1	YCBL_ECOLI	HYPOTHETICAL 23.8 KD P	2.22e+03
304	4	13.8	147	1	YBAK_BACSU	HYPOTHETICAL 17.6 KD P	2.22e+03	377	4	13.8	216	1	YGD7_YEAST	HYPOTHETICAL 25.0 KD P	2.22e+03
305	4	13.8	147	1	Y226_MYCTU	HYPOTHETICAL 16.4 KD P	2.22e+03	378	4	13.8	217	1	YAAE_BACSU	HYPOTHETICAL 25.4 KD P	2.22e+03
306	4	13.8	147	1	Y220_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	379	4	13.8	218	1	Y232_SYNY3	HYPOTHETICAL 23.8 KD P	2.22e+03
307	4	13.8	148	1	YLL2_EBVA8	HYPOTHETICAL BLLF2 PRO	2.22e+03	380	4	13.8	219	1	Y364_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
308	4	13.8	148	1	YLL2_EBV	HYPOTHETICAL BLLF2 PRO	2.22e+03	381	4	13.8	219	1	Y210_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
309	4	13.8	149	1	YML3_THIFE	HYPOTHETICAL 15.6 KD P	2.22e+03	382	4	13.8	221	1	YPRB_CORGL	HYPOTHETICAL PROTEIN I	2.22e+03
310	4	13.8	153	1	YS12_CABEL	HYPOTHETICAL 17.8 KD P	2.22e+03	383	4	13.8	225	1	Y06L_MYCTU	PROBABLE POLYKETIDE SY	2.22e+03
311	4	13.8	153	1	Y01J_MYCTU	HYPOTHETICAL 16.8 KD P	2.22e+03	384	4	13.8	226	1	VP09_BPRD	DNA PACKAGING PROTEIN	2.22e+03
312	4	13.8	153	1	YPG_BYDVP	PUTATIVE GENOME-LINKED	2.22e+03	385	4	13.8	227	1	Y295_METJA	HYPOTHETICAL PROTEIN H	2.22e+03
313	4	13.8	153	1	Y040_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03	386	4	13.8	227	1	YD31_HELPY	HYPOTHETICAL PROTEIN H	2.22e+03
314	4	13.8	154	1	YWEA_BACSU	HYPOTHETICAL 16.7 KD P	2.22e+03	387	4	13.8	229	1	YXHK_CABEL	HYPOTHETICAL 25.4 KD P	2.22e+03
315	4	13.8	156	1	VSR_ECOLI	PATCH REPAIR PROTEIN (	2.22e+03	388	4	13.8	234	1	VPRT_HTLIA	PROTEASE (EC 3.4.23.-)	2.22e+03

389	4	13.8	234	1	Y340_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
390	4	13.8	236	1	YP2C_STAAU	HYPOTHETICAL 27.7 KD P	2.22e+03
391	4	13.8	238	1	YP22_RHORU	HYPOTHETICAL 25.8 KD P	2.22e+03
392	4	13.8	240	1	YX22_SCHPO	HYPOTHETICAL 28.1 KD P	2.22e+03
393	4	13.8	240	1	YDPE_SCHPO	HYPOTHETICAL 27.3 KD P	2.22e+03
394	4	13.8	240	1	Y354_HAEIN	HYPOTHETICAL ABC TRANS	2.22e+03
395	4	13.8	240	1	ZE7_MAIZE	ZEIN-ALPHA PRECURSOR (	2.22e+03
396	4	13.8	240	1	YQ12_BACSU	PROBABLE AMINO-ACID AB	2.22e+03
397	4	13.8	240	1	ZE9_MAIZE	ZEIN-ALPHA PRECURSOR (	2.22e+03
398	4	13.8	240	1	ZE6_MAIZE	ZEIN-ALPHA PRECURSOR (	2.22e+03
399	4	13.8	240	1	ZE8_MAIZE	ZEIN-ALPHA PRECURSOR (	2.22e+03
400	4	13.8	243	1	YQF6_BACSU	HYPOTHETICAL 27.6 KD L	2.22e+03
401	4	13.8	245	1	YJF5_YEAST	HYPOTHETICAL 26.9 KD P	2.22e+03
402	4	13.8	245	1	YAB5_BACSU	HYPOTHETICAL 26.8 KD P	2.22e+03
403	4	13.8	247	1	YJC1_ECOLI	HYPOTHETICAL 26.4 KD P	2.22e+03
404	4	13.8	247	1	YABB_BACSU	HYPOTHETICAL 28.3 KD P	2.22e+03
405	4	13.8	249	1	YJFP_ECOLI	HYPOTHETICAL 27.6 KD P	2.22e+03
406	4	13.8	250	1	YL22_YEAST	HYPOTHETICAL 28.3 KD P	2.22e+03
407	4	13.8	250	1	YJ05_YEAST	HYPOTHETICAL 28.3 KD P	2.22e+03
408	4	13.8	251	1	Y545_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
409	4	13.8	251	1	YJF0_ECOLI	HYPOTHETICAL TRANSCRIP	2.22e+03
410	4	13.8	251	1	YH64_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
411	4	13.8	252	1	YH02_ECOLI	HYPOTHETICAL AMINO-ACI	2.22e+03
412	4	13.8	252	1	YJG6_YEAST	HYPOTHETICAL 28.5 KD P	2.22e+03
413	4	13.8	254	1	YK07_CAEEL	HYPOTHETICAL 29.5 KD P	2.22e+03
414	4	13.8	254	1	YD1J_BACSU	HYPOTHETICAL 29.1 KD P	2.22e+03
415	4	13.8	254	1	YCAL_ECOLI	HYPOTHETICAL 26.7 KD P	2.22e+03
416	4	13.8	254	1	YCA5_YEAST	HYPOTHETICAL 28.7 KD P	2.22e+03
417	4	13.8	255	1	YW24_MYCTU	HYPOTHETICAL 28.1 KD P	2.22e+03
418	4	13.8	258	1	YDF2_SCHPO	HYPOTHETICAL 29.3 KD P	2.22e+03
419	4	13.8	258	1	YAAA_ECOLI	HYPOTHETICAL 29.6 KD P	2.22e+03
420	4	13.8	259	1	YGA7_SYNY3	HYPOTHETICAL 28.1 KD P	2.22e+03
421	4	13.8	259	1	YAO1_SCHPO	HYPOTHETICAL 30.0 KD P	2.22e+03
422	4	13.8	260	1	YQCB_ECOLI	HYPOTHETICAL 29.7 KD P	2.22e+03
423	4	13.8	261	1	YG08_YEAST	HYPOTHETICAL 29.1 KD P	2.22e+03
424	4	13.8	261	1	Y09T_MYCLE	HYPOTHETICAL 28.6 KD P	2.22e+03
425	4	13.8	261	1	YQST_VERPS	YOP PROTEINS TRANSLOCA	2.22e+03
426	4	13.8	262	1	YQ52_CAEEL	HYPOTHETICAL 30.9 KD P	2.22e+03
427	4	13.8	263	1	Y314_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
428	4	13.8	263	1	YCA3_OOOSI	HYPOTHETICAL 30.1 KD P	2.22e+03
429	4	13.8	263	1	YBGF_ECOLI	HYPOTHETICAL 28.2 KD P	2.22e+03
430	4	13.8	264	1	YHFJ_SALTU	HYPOTHETICAL 26.3 KD P	2.22e+03
431	4	13.8	265	1	YEB1_ECOLI	HYPOTHETICAL 30.3 KD P	2.22e+03
432	4	13.8	266	1	YD37_SCHPO	VERY HYPOTHETICAL 31.0	2.22e+03
433	4	13.8	268	1	YL12_MYCHO	HYPOTHETICAL 31.2 KD P	2.22e+03
434	4	13.8	268	1	Y145_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
435	4	13.8	269	1	YDUD_SCHPO	HYPOTHETICAL 29.9 KD P	2.22e+03
436	4	13.8	269	1	YBBO_ECOLI	HYPOTHETICAL OXIDOREDU	2.22e+03
437	4	13.8	270	1	YNF0_YEAST	HYPOTHETICAL 31.4 KD P	2.22e+03
438	4	13.8	271	1	Y398_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
439	4	13.8	271	1	Y01G_MYCTU	HYPOTHETICAL 29.8 KD P	2.22e+03
440	4	13.8	272	1	Y4YN_RHISN	PROBABLE TRANSLOCATION	2.22e+03
441	4	13.8	273	1	Y7Y9_SCHPO	HYPOTHETICAL 32.5 KD P	2.22e+03
442	4	13.8	273	1	Y030_MYCTU	HYPOTHETICAL 29.2 KD P	2.22e+03
443	4	13.8	274	1	Y179_MYCGE	HYPOTHETICAL ABC TRANS	2.22e+03
444	4	13.8	274	1	Y179_MYCPN	HYPOTHETICAL ABC TRANS	2.22e+03
445	4	13.8	274	1	YOHM_ECOLI	HYPOTHETICAL 30.4 KD P	2.22e+03
446	4	13.8	275	1	Y31K_SULAC	HYPOTHETICAL 31.5 KD P	2.22e+03
447	4	13.8	277	1	Y4FA_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03
448	4	13.8	278	1	YMP5_STRCO	HYPOTHETICAL PROTEIN I	2.22e+03
449	4	13.8	278	1	YQ66_CAEEL	HYPOTHETICAL 32.5 KD P	2.22e+03
450	4	13.8	278	1	Y61G_ECOLI	HYPOTHETICAL 31.3 KD P	2.22e+03
451	4	13.8	279	1	YKDY_BACSU	PHAGE-LIKE ELEMENT PBS	2.22e+03
452	4	13.8	279	1	YS11_CAEEL	HYPOTHETICAL 31.2 KD P	2.22e+03
453	4	13.8	281	1	YBXA_BACSU	HYPOTHETICAL ABC TRANS	2.22e+03
454	4	13.8	282	1	YH0R_ECOLI	HYPOTHETICAL 31.5 KD P	2.22e+03
455	4	13.8	282	1	Y265_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
456	4	13.8	284	1	Y030_MYCLE	HYPOTHETICAL 30.3 KD P	2.22e+03
457	4	13.8	286	1	Y320_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
458	4	13.8	290	1	YX17_MYCTU	HYPOTHETICAL 33.0 KD P	2.22e+03
459	4	13.8	291	1	YC43_GUITH	HYPOTHETICAL 33.1 KD P	2.22e+03
460	4	13.8	291	1	Y32K_SSV1	HYPOTHETICAL 31.7 KD P	2.22e+03
461	4	13.8	291	1	YG47_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
462	4	13.8	291	1	Y620_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
463	4	13.8	291	1	YJ15_PYRHO	HYPOTHETICAL ABC TRANS	2.22e+03
464	4	13.8	292	1	YXG_SCHPO	HYPOTHETICAL PROTEIN C	2.22e+03
465	4	13.8	292	1	YPRS_CHLTR	HYPOTHETICAL 31.6 KD P	2.22e+03
466	4	13.8	293	1	YQFU_BACSU	HYPOTHETICAL 32.5 KD P	2.22e+03
467	4	13.8	293	1	Y1A3_YEAST	HYPOTHETICAL 31.9 KD P	2.22e+03
468	4	13.8	294	1	YAFJ_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
469	4	13.8	294	1	YX23_MAIZE	HYPOTHETICAL 33.9 KD P	2.22e+03
470	4	13.8	295	1	YCX7_EUGGR	HYPOTHETICAL 35.4 KD P	2.22e+03
471	4	13.8	296	1	YWFJ_BACSU	HYPOTHETICAL 31.3 KD P	2.22e+03
472	4	13.8	297	1	YVPR_SALDU	VIRULENCE GENES TRANS	2.22e+03
473	4	13.8	297	1	YCD1_YEAST	HYPOTHETICAL 34.5 KD P	2.22e+03
474	4	13.8	297	1	YC24_ANTSP	HYPOTHETICAL PROTEIN I	2.22e+03
475	4	13.8	298	1	YNNB_ECOLI	HYPOTHETICAL 33.1 KD P	2.22e+03
476	4	13.8	298	1	YNN5_YEAST	HYPOTHETICAL 34.5 KD P	2.22e+03
477	4	13.8	298	1	Y664_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
478	4	13.8	301	1	YWB1_BACSU	HYPOTHETICAL TRANSCRIP	2.22e+03
479	4	13.8	301	1	YEQ6_YEAST	HYPOTHETICAL 33.5 KD P	2.22e+03
480	4	13.8	302	1	YXNA_THEAU	ENDO-1,4-BETA-XYLANASE	2.22e+03
481	4	13.8	302	1	Y091_METJA	HYPOTHETICAL 35.5 KD P	2.22e+03
482	4	13.8	302	1	YD13_SCHPO	ENDO-1,4-BETA-XYLANASE	2.22e+03
483	4	13.8	302	1	YXNA_PENSI	HYPOTHETICAL 35.5 KD P	2.22e+03
484	4	13.8	302	1	YXID_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
485	4	13.8	303	1	Y180_MYCPN	HYPOTHETICAL ABC TRANS	2.22e+03
486	4	13.8	305	1	YQ31_STRGR	HYPOTHETICAL 31.2 KD P	2.22e+03
487	4	13.8	306	1	YQKF_BACSU	HYPOTHETICAL OXIDOREDU	2.22e+03
488	4	13.8	306	1	YQ23_HCMVA	PROBABLE CAPSID PROTEI	2.22e+03
489	4	13.8	307	1	YXET_PSEPU	METAPYROCATECHASE (EC	2.22e+03
490	4	13.8	307	1	YXEL_PSEAE	METAPYROCATECHASE (EC	2.22e+03
491	4	13.8	307	1	YXEL_PSEAE	METAPYROCATECHASE (EC	2.22e+03
492	4	13.8	308	1	YNEH_ECOLI	HYPOTHETICAL 33.5 KD P	2.22e+03
493	4	13.8	308	1	YBIN_ECOLI	HYPOTHETICAL 34.2 KD P	2.22e+03
494	4	13.8	308	1	YQAB_SCHPO	HYPOTHETICAL 35.3 KD P	2.22e+03
495	4	13.8	310	1	Y094_METJA	HYPOTHETICAL ABC TRANS	2.22e+03
496	4	13.8	311	1	Y467_MYCGE	HYPOTHETICAL 34.1 KD P	2.22e+03
497	4	13.8	312	1	YKXC_BACSU	PROBABLE PEPTIDE ABC T	2.22e+03
498	4	13.8	313	1	Y4TP_RHISN	PROBABLE PEPTIDE ABC T	2.22e+03
499	4	13.8	313	1	Y4JF_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
500	4	13.8	313	1	Y4JF_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
501	4	13.8	314	1	YHFM_HAEIN	LYTB PROTEIN HOMOLOG.	2.22e+03
502	4	13.8	314	1	YQFP_BACSU	HYPOTHETICAL 35.9 KD P	2.22e+03
503	4	13.8	314	1	YC39_GUITH	HYPOTHETICAL 35.9 KD P	2.22e+03
504	4	13.8	315	1	YK04_CAEEL	Y BOX BINDING PROTEIN-	2.22e+03
505	4	13.8	317	1	YB1_HUMAN	Y BOX BINDING PROTEIN-	2.22e+03
506	4	13.8	318	1	Y1L7_YEAST	HYPOTHETICAL 34.7 KD P	2.22e+03
507	4	13.8	319	1	YQ39_PORPU	HYPOTHETICAL 35.7 KD P	2.22e+03
508	4	13.8	320	1	YQ19_KLEPN	HYPOTHETICAL PROTEIN I	2.22e+03
509	4	13.8	320	1	Y678_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
510	4	13.8	321	1	YB1_CHICK	Y BOX BINDING PROTEIN-	2.22e+03
511	4	13.8	321	1	YHBE_ECOLI	Y BOX BINDING PROTEIN-	2.22e+03
512	4	13.8	322	1	YB1_MOUSE	Y BOX BINDING PROTEIN-	2.22e+03
513	4	13.8	322	1	Y1L2_SYNY3	HYPOTHETICAL 35.1 KD P	2.22e+03
514	4	13.8	322	1	Y048_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
515	4	13.8	322	1	Y4UB_SCHPO	HYPOTHETICAL 35.7 KD P	2.22e+03
516	4	13.8	323	1	Y370_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
517	4	13.8	323	1	Y1V6_MYCTU	HYPOTHETICAL 33.3 KD P	2.22e+03
518	4	13.8	325	1	VP6_BT113	VP6 PROTEIN (MINOR INN	2.22e+03
519	4	13.8	325	1	VP6_BT111	VP6 PROTEIN (MINOR INN	2.22e+03
520	4	13.8	326	1	Y370_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
521	4	13.8	326	1	YR75_ECOLI	HYPOTHETICAL 36.8 KD P	2.22e+03
522	4	13.8	326	1	YF02_MYCPN	HYPOTHETICAL ABC TRANS	2.22e+03
523	4	13.8	327	1	YXNA_ASPAK	ENDO-1,4-BETA-XYLANASE	2.22e+03
524	4	13.8	327	1	YX91_YEAST	HYPOTHETICAL 35.7 KD P	2.22e+03
525	4	13.8	329	1	YX88_YEAST	HYPOTHETICAL 37.1 KD P	2.22e+03
526	4	13.8	329	1	Y1V9_MYCTU	HYPOTHETICAL 35.4 KD P	2.22e+03
527	4	13.8	329	1	YTXK_BACSU	HYPOTHETICAL 37.4 KD P	2.22e+03
528	4	13.8	329	1	YF61_BT110	VP6 PROTEIN (MINOR INN	2.22e+03
529	4	13.8	330	1	Y367_METJA	PROBABLE INTEGRASE/REC	2.22e+03
530	4	13.8	330	1	Y273_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
531	4	13.8	331	1	VP4_BPPH5	P4 PROTEIN	2.22e+03
532	4	13.8	334	1	YF41_ARATH	VACUOLAR ASSEMBLY PROT	2.22e+03
533	4	13.8	335	1	Y1L2_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
534	4	13.8	335	1	YEFK_ECOLI	HYPOTHETICAL 37.8 KD P	2.22e+03



535	4	13.8	336	1	YD9G_SCHPO	HYPOTHETICAL PROTEIN C	2.22e+03	608	4	13.8	392	1	YQF6_CAEEL	HYPOTHETICAL 45.3 KD P	2.22e+03
536	4	13.8	336	1	Y0SG_BPT4	HYPOTHETICAL 38.9 KD P	2.22e+03	609	4	13.8	392	1	Y4FQ_RHISN	HYPOTHETICAL 41.6 KD P	2.22e+03
537	4	13.8	337	1	YHUD_ECOLI	HYPOTHETICAL 37.9 KD P	2.22e+03	610	4	13.8	393	1	Y306_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
538	4	13.8	337	1	ZNUA_HAEIN	HIGH-AFFINITY ZINC UPT	2.22e+03	611	4	13.8	393	1	YCFU_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
539	4	13.8	339	1	YEA9_SCHPO	HYPOTHETICAL 38.5 KD P	2.22e+03	612	4	13.8	393	1	XYLH_ECOLI	XYLOSE TRANSPORT PERME	2.22e+03
540	4	13.8	339	1	Y467_MYCPN	HYPOTHETICAL ABC TRANS	2.22e+03	613	4	13.8	399	1	Y017_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
541	4	13.8	339	1	Y090_SYNY3	HYPOTHETICAL 38.6 KD P	2.22e+03	614	4	13.8	401	1	YABA_SCHPO	HYPOTHETICAL 44.4 KD P	2.22e+03
542	4	13.8	340	1	YMD7_YEAST	HYPOTHETICAL 38.0 KD P	2.22e+03	615	4	13.8	401	1	Y148_METSM	INSERTION ELEMENT ISM1	2.22e+03
543	4	13.8	340	1	Y269_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03	616	4	13.8	401	1	Y766_CAEEL	PROBABLE G PROTEIN-COU	2.22e+03
544	4	13.8	341	1	YHS3_SACKL	HYPOTHETICAL PROTEIN I	2.22e+03	617	4	13.8	402	1	Y718_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
545	4	13.8	341	1	YB57_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	618	4	13.8	403	1	YBL2_SFV1	BEL-2 PROTEIN.	2.22e+03
546	4	13.8	341	1	YD1F_SCHPO	HYPOTHETICAL OXIDOREDU	2.22e+03	619	4	13.8	403	1	YLAO_BACSU	HYPOTHETICAL 43.7 KD P	2.22e+03
547	4	13.8	341	1	Y665_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	620	4	13.8	403	1	YGEY_ECOLI	HYPOTHETICAL 44.8 KD P	2.22e+03
548	4	13.8	342	1	YJ8K_ECOLI	HYPOTHETICAL 39.7 KD P	2.22e+03	621	4	13.8	404	1	Y4XM_RHISN	HYPOTHETICAL TRANSPORT	2.22e+03
549	4	13.8	342	1	YH06_YEAST	HYPOTHETICAL 37.1 KD P	2.22e+03	622	4	13.8	404	1	YBR3_YEAST	HYPOTHETICAL 45.4 KD P	2.22e+03
550	4	13.8	345	1	YCX1_CHLMO	HYPOTHETICAL 38.5 KD P	2.22e+03	623	4	13.8	405	1	YEAH_SCHPO	HYPOTHETICAL 45.6 KD P	2.22e+03
551	4	13.8	345	1	YGV3_YEAST	HYPOTHETICAL 39.6 KD P	2.22e+03	624	4	13.8	406	1	YBDN_ECOLI	HYPOTHETICAL 47.8 KD P	2.22e+03
552	4	13.8	345	1	YD59_SCHPO	PUTATIVE MITOCHONDRIAL	2.22e+03	625	4	13.8	406	1	Y129_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
553	4	13.8	346	1	YD6J_ECOLI	HYPOTHETICAL OXIDOREDU	2.22e+03	626	4	13.8	408	1	YCEE_ECOLI	HYPOTHETICAL 43.9 KD P	2.22e+03
554	4	13.8	346	1	YRAQ_ECOLI	HYPOTHETICAL 37.3 KD P	2.22e+03	627	4	13.8	408	1	Y899_SYNY3	HYPOTHETICAL 46.0 KD P	2.22e+03
555	4	13.8	346	1	YIC3_YEAST	HYPOTHETICAL 37.4 KD P	2.22e+03	628	4	13.8	411	1	YPEB_SYNP2	HYPOTHETICAL PROTEIN I	2.22e+03
556	4	13.8	346	1	Y285_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03	629	4	13.8	411	1	XYNA_BUTFI	ENDO-1,4-BETA-XYLANASE	2.22e+03
557	4	13.8	347	1	YQ9N_BACSU	HYPOTHETICAL 40.5 KD P	2.22e+03	630	4	13.8	412	1	YBDR_ECOLI	HYPOTHETICAL ZINC-TYPE	2.22e+03
558	4	13.8	349	1	YLBC_ECOLI	HYPOTHETICAL 38.0 KD P	2.22e+03	631	4	13.8	413	1	YAA7_SCHPO	HYPOTHETICAL 47.3 KD P	2.22e+03
559	4	13.8	350	1	YURN_BACSU	HYPOTHETICAL LIPOPROTE	2.22e+03	632	4	13.8	414	1	Y288_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
560	4	13.8	351	1	YIMZ_YEAST	HYPOTHETICAL 39.5 KD P	2.22e+03	633	4	13.8	414	1	YM87_YEAST	HYPOTHETICAL 46.7 KD P	2.22e+03
561	4	13.8	351	1	YE72_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	634	4	13.8	416	1	YMBN_BACSU	HYPOTHETICAL 45.7 KD P	2.22e+03
562	4	13.8	352	1	YFEA_ECOLI	HYPOTHETICAL 40.0 KD P	2.22e+03	635	4	13.8	419	1	YDEA_SCHPO	HYPOTHETICAL 48.1 KD P	2.22e+03
563	4	13.8	352	1	YDGO_ECOLI	HYPOTHETICAL 38.1 KD P	2.22e+03	636	4	13.8	421	1	YK18_YEAST	HYPOTHETICAL 46.6 KD P	2.22e+03
564	4	13.8	353	1	ZIC4_MOUSE	ZINC FINGER PROTEIN ZI	2.22e+03	637	4	13.8	421	1	YEO4_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
565	4	13.8	354	1	YGP1_YEAST	PROTEIN YGP1 PRECURSOR	2.22e+03	638	4	13.8	422	1	YUS3_YEAST	HYPOTHETICAL 47.8 KD P	2.22e+03
566	4	13.8	354	1	YQ11_ECOLI	HYPOTHETICAL 38.6 KD P	2.22e+03	639	4	13.8	422	1	YUHO_BACSU	HYPOTHETICAL ABC TRANS	2.22e+03
567	4	13.8	354	1	YSCU_YERPS	YOP PROTEINS TRANSLOCA	2.22e+03	640	4	13.8	423	1	YDAJ_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03
568	4	13.8	355	1	YM79_YEAST	HYPOTHETICAL 37.4 KD P	2.22e+03	641	4	13.8	423	1	YHJ3_YEAST	HYPOTHETICAL 46.9 KD P	2.22e+03
569	4	13.8	357	1	YE38_HELPY	HYPOTHETICAL PROTEIN H	2.22e+03	642	4	13.8	424	1	Y146_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
570	4	13.8	358	1	Y4EE_RHISN	HYPOTHETICAL 38.5 KD P	2.22e+03	643	4	13.8	424	1	VP3_BPHK7	PORTAL PROTEIN (GP3)	2.22e+03
571	4	13.8	359	1	YKAT_CAEEL	HYPOTHETICAL 39.7 KD P	2.22e+03	644	4	13.8	425	1	YEGT_ECOLI	PUTATIVE NUCLEOSIDE TR	2.22e+03
572	4	13.8	359	1	YQ8M_BACSU	HYPOTHETICAL 40.8 KD P	2.22e+03	645	4	13.8	425	1	YH84_CAEEL	PROBABLE G PROTEIN-COU	2.22e+03
573	4	13.8	360	1	YACK_BACSU	HYPOTHETICAL 40.7 KD P	2.22e+03	646	4	13.8	425	1	YU77_YEAST	HYPOTHETICAL 49.1 KD P	2.22e+03
574	4	13.8	360	1	Y33_ARMV	P3 PROTEIN.	2.22e+03	647	4	13.8	425	1	YTR1_EBV	HYPOTHETICAL BTRFI PRO	2.22e+03
575	4	13.8	361	1	YC25_YEAST	HYPOTHETICAL ZINC-TYPE	2.22e+03	648	4	13.8	425	1	Y461_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
576	4	13.8	361	1	YB1C_ECOLI	HYPOTHETICAL 38.9 KD P	2.22e+03	649	4	13.8	427	1	YIEM_ECOLI	HYPOTHETICAL 49.6 KD P	2.22e+03
577	4	13.8	363	1	YRAK_ECOLI	HYPOTHETICAL 38.5 KD P	2.22e+03	650	4	13.8	427	1	YBBS_ECOLI	HYPOTHETICAL 48.3 KD P	2.22e+03
578	4	13.8	363	1	YQ42_CAEEL	HYPOTHETICAL 40.0 KD P	2.22e+03	651	4	13.8	427	1	Y96E_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
579	4	13.8	364	1	YBL2_FOAMV	BEL-2 PROTEIN.	2.22e+03	652	4	13.8	431	1	Y4OP_RHISN	PROBABLE AMINO-ACID AB	2.22e+03
580	4	13.8	364	1	YH1M_ECOLI	HYPOTHETICAL 39.2 KD P	2.22e+03	653	4	13.8	432	1	YNN3_PODAN	HYPOTHETICAL 49.1 KD P	2.22e+03
581	4	13.8	364	1	Y174_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03	654	4	13.8	433	1	YTFL_HAEIN	HYPOTHETICAL 49.1 KD P	2.22e+03
582	4	13.8	366	1	YNR3_YEAST	HYPOTHETICAL 40.3 KD P	2.22e+03	655	4	13.8	433	1	YGCN_ECOLI	PROBABLE ELECTRON TRAN	2.22e+03
583	4	13.8	366	1	YRAF_BACSU	HYPOTHETICAL 40.1 KD G	2.22e+03	656	4	13.8	434	1	YRL4_CAEEL	HYPOTHETICAL 49.2 KD P	2.22e+03
584	4	13.8	367	1	YV29_MYCLE	HYPOTHETICAL 40.5 KD P	2.22e+03	657	4	13.8	434	1	YAOA_SCHPO	HYPOTHETICAL 48.2 KD P	2.22e+03
585	4	13.8	368	1	Y974_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	658	4	13.8	435	1	Y609_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
586	4	13.8	368	1	YC42_YEAST	40.1 KD PROTEIN IN GIT	2.22e+03	659	4	13.8	436	1	Y326_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
587	4	13.8	372	1	Y412_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03	660	4	13.8	439	1	Y204_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
588	4	13.8	372	1	YJGP_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	661	4	13.8	440	1	YHR1_YEAST	HYPOTHETICAL 49.4 KD P	2.22e+03
589	4	13.8	373	1	YF34_YEAST	HYPOTHETICAL 40.4 KD P	2.22e+03	662	4	13.8	440	1	YMI9_MYCTU	HYPOTHETICAL 45.3 KD T	2.22e+03
590	4	13.8	373	1	YWAP_STRMU	HYPOTHETICAL PROTEIN I	2.22e+03	663	4	13.8	442	1	ZRCL_YEAST	ZINC/CADMIUM RESISTANC	2.22e+03
591	4	13.8	374	1	YLD1_CAEEL	PROBABLE G PROTEIN-COU	2.22e+03	664	4	13.8	443	1	YR15_CAEEL	HYPOTHETICAL 49.8 KD P	2.22e+03
592	4	13.8	375	1	Y4ED_RHISN	HYPOTHETICAL 42.2 KD P	2.22e+03	665	4	13.8	443	1	Y314_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
593	4	13.8	376	1	YL60_YEAST	HYPOTHETICAL 41.1 KD P	2.22e+03	666	4	13.8	443	1	YAAU_ECOLI	HYPOTHETICAL METABOLIT	2.22e+03
594	4	13.8	376	1	Y023_BPHPI	PUTATIVE TAIL SHEATH P	2.22e+03	667	4	13.8	444	1	YOR3_GLSU	HYPOTHETICAL PROTEIN I	2.22e+03
595	4	13.8	376	1	ZRT1_PROTEIN.	ZRT1 PROTEIN.	2.22e+03	668	4	13.8	445	1	YGC5_ECOLI	HYPOTHETICAL METABOLIT	2.22e+03
596	4	13.8	378	1	YD40_YEAST	HYPOTHETICAL 42.9 KD P	2.22e+03	669	4	13.8	445	1	YIEG_ECOLI	HYPOTHETICAL 46.9 KD P	2.22e+03
597	4	13.8	378	1	YMX5_CAEEL	HYPOTHETICAL 43.3 KD P	2.22e+03	670	4	13.8	446	1	YKAI_COXBU	HYPOTHETICAL 49.9 KD P	2.22e+03
598	4	13.8	378	1	YNR8_YEAST	HYPOTHETICAL 41.9 KD P	2.22e+03	671	4	13.8	447	1	YKCT_YEAST	HYPOTHETICAL 50.3 KD P	2.22e+03
599	4	13.8	379	1	YLEU_DEBOC	HYPOTHETICAL 43.6 KD P	2.22e+03	672	4	13.8	447	1	YAWB_RHISN	HYPOTHETICAL ZINC PROT	2.22e+03
600	4	13.8	379	1	Y024_BORBU	HYPOTHETICAL PROTEIN B	2.22e+03	673	4	13.8	448	1	YJ94_YEAST	HYPOTHETICAL 49.7 KD P	2.22e+03
601	4	13.8	381	1	YNU6_CAEEL	HYPOTHETICAL 42.9 KD P	2.22e+03	674	4	13.8	448	1	Y260_SYNY3	HYPOTHETICAL 49.2 KD P	2.22e+03
602	4	13.8	381	1	YS09_MYCTU	HYPOTHETICAL 41.2 KD P	2.22e+03	675	4	13.8	448	1	YAE5_CHICK	65 KD YES-ASSOCIATED P	2.22e+03
603	4	13.8	385	1	YF00_YEAST	HYPOTHETICAL 41.9 KD P	2.22e+03	676	4	13.8	449	1	XYLA_LACBR	XYLOSE ISOMERASE (EC 5	2.22e+03
604	4	13.8	386	1	YGV5_YEAST	HYPOTHETICAL 40.2 KD P	2.22e+03	677	4	13.8	449	1	YGIY_ECOLI	PROBABLE SENSOR PROTEI	2.22e+03
605	4	13.8	387	1	YHFJ_ECOLI	HYPOTHETICAL 44.3 KD P	2.22e+03	678	4	13.8	450	1	YK22_YEAST	HYPOTHETICAL 46.9 KD P	2.22e+03
606	4	13.8	388	1	YFHS_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	679	4	13.8	451	1	Y4T1_RHISN	PROBABLE ATP SYNTHASE	2.22e+03
607	4	13.8	390	1	Y4R1_RHISN	HYPOTHETICAL 44.0 KD P	2.22e+03	680	4	13.8	451	1	YGIY_HAEIN	PROBABLE SENSOR PROTEI	2.22e+03

681	4	13.8	452	1	YDQ4_SCHPO	HYPOTHETICAL 49.6 KD P	2.22e+03	754	4	13.8	544	1	YL33_CAEEL	HYPOTHETICAL 61.8 KD P	2.22e+03
682	4	13.8	452	1	YX37_MYCTU	HYPOTHETICAL 51.4 KD P	2.22e+03	755	4	13.8	544	1	YT25_ARATH	HYPOTHETICAL 62.3 KD P	2.22e+03
683	4	13.8	454	1	YAG5_HUMAN	65 KD YES-ASSOCIATED P	2.22e+03	756	4	13.8	546	1	Y772_SYNV3	HYPOTHETICAL 59.1 KD P	2.22e+03
684	4	13.8	455	1	YICL_ECOLI	HYPOTHETICAL 48.8 KD P	2.22e+03	757	4	13.8	547	1	YOJ1_ECOLI	HYPOTHETICAL ABC TRANS	2.22e+03
685	4	13.8	456	1	YEB8_SCHPO	HYPOTHETICAL 52.4 KD P	2.22e+03	758	4	13.8	549	1	YJCE_ECOLI	HYPOTHETICAL 60.5 KD P	2.22e+03
686	4	13.8	457	1	YICB_ECOLI	HYPOTHETICAL 50.3 KD L	2.22e+03	759	4	13.8	552	1	YF64_HAEIN	HYPOTHETICAL ABC TRANS	2.22e+03
687	4	13.8	458	1	Y834_HELPY	HYPOTHETICAL GTP-BINDI	2.22e+03	760	4	13.8	553	1	YF61_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
688	4	13.8	461	1	Y061_MYCGE	HYPOTHETICAL 50.3 KD P	2.22e+03	761	4	13.8	554	1	Y514_SYNV3	HYPOTHETICAL 62.4 KD P	2.22e+03
689	4	13.8	462	1	Y036_YEAST	HYPOTHETICAL PROTEIN M	2.22e+03	762	4	13.8	558	1	Y322_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
690	4	13.8	463	1	Y05F_MYCTU	HYPOTHETICAL 53.4 KD P	2.22e+03	763	4	13.8	558	1	YCX9_ASILLO	HYPOTHETICAL 66.5 KD P	2.22e+03
691	4	13.8	464	1	YEM8_YEAST	HYPOTHETICAL 46.9 KD P	2.22e+03	764	4	13.8	561	1	YGG4_YEAST	PUTATIVE ATP-DEPENDENT	2.22e+03
692	4	13.8	464	1	Y098_YEAST	HYPOTHETICAL 53.9 KD P	2.22e+03	765	4	13.8	563	1	YIMO_YEAST	HYPOTHETICAL 61.8 KD P	2.22e+03
693	4	13.8	465	1	YIC4_KLEPN	PROBABLE CPS BIOSYNTH	2.22e+03	766	4	13.8	563	1	Y322_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
694	4	13.8	465	1	Y957_YEAST	HYPOTHETICAL 53.1 KD T	2.22e+03	767	4	13.8	566	1	YOK4_CAEEL	HYPOTHETICAL 64.3 KD P	2.22e+03
695	4	13.8	466	1	Y495_YEAST	HYPOTHETICAL GTP-BINDI	2.22e+03	768	4	13.8	571	1	YHY4_YEAST	HYPOTHETICAL 65.8 KD P	2.22e+03
696	4	13.8	471	1	YODO_BACSU	HYPOTHETICAL 54.1 KD P	2.22e+03	769	4	13.8	573	1	YEV4_YEAST	HYPOTHETICAL 64.0 KD P	2.22e+03
697	4	13.8	471	1	Y782_SYNV3	HYPOTHETICAL 51.4 KD P	2.22e+03	770	4	13.8	573	1	YFIB_BACSU	HYPOTHETICAL ABC TRANS	2.22e+03
698	4	13.8	475	1	YM61_YEAST	HYPOTHETICAL 55.3 KD P	2.22e+03	771	4	13.8	575	1	YWJA_BACSU	HYPOTHETICAL ABC TRANS	2.22e+03
699	4	13.8	476	1	Y554_CAEEL	HYPOTHETICAL 53.9 KD P	2.22e+03	772	4	13.8	576	1	Y134_HUMAN	PUTATIVE ATP-DEPENDENT	2.22e+03
700	4	13.8	477	1	Y098_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03	773	4	13.8	577	1	YG5U_YEAST	HYPOTHETICAL 67.3 KD P	2.22e+03
701	4	13.8	477	1	Y100_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03	774	4	13.8	577	1	YV45_YEAST	VACUOLAR PROTEIN SORTI	2.22e+03
702	4	13.8	478	1	YSP3_YEAST	SUBILISIN-LIKE PROTEA	2.22e+03	775	4	13.8	577	1	YTFM_ECOLI	HYPOTHETICAL 64.8 KD P	2.22e+03
703	4	13.8	479	1	YBHI_HAEIN	HYPOTHETICAL PROTEIN H	2.22e+03	776	4	13.8	581	1	YHJ2_YEAST	HYPOTHETICAL 64.2 KD P	2.22e+03
704	4	13.8	480	1	Y99Q_YEAST	HYPOTHETICAL 55.5 KD P	2.22e+03	777	4	13.8	585	1	YB13_YEAST	HYPOTHETICAL 67.6 KD P	2.22e+03
705	4	13.8	482	1	Y138_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	778	4	13.8	585	1	YHD9_YEAST	HYPOTHETICAL 67.5 KD P	2.22e+03
706	4	13.8	485	1	Z165_HUMAN	ZINC FINGER PROTEIN 16	2.22e+03	779	4	13.8	586	1	YV40_ILTVT	CAPSID PROTEIN P40 [CO	2.22e+03
707	4	13.8	488	1	ZF92_MOUSE	ZINC FINGER PROTEIN 92	2.22e+03	780	4	13.8	586	1	Y4GM_RHEIN	PROBABLE ABC TRANSPORT	2.22e+03
708	4	13.8	489	1	YU42_MYCTU	PUTATIVE CYTOCHROME P4	2.22e+03	781	4	13.8	587	1	YID6_YEAST	HYPOTHETICAL 65.3 KD P	2.22e+03
709	4	13.8	491	1	YF86_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	782	4	13.8	589	1	YAN6_SCHPO	HYPOTHETICAL 63.5 KD P	2.22e+03
710	4	13.8	492	1	YEO1_YEAST	HYPOTHETICAL 56.5 KD P	2.22e+03	783	4	13.8	592	1	Y036_HAEIN	HYPOTHETICAL ABC TRANS	2.22e+03
711	4	13.8	492	1	YSM1_TRYBB	VARIANT SURFACE GLYCOP	2.22e+03	784	4	13.8	593	1	Y1BB_MYCTU	PROBABLE TRANSPORT PRO	2.22e+03
712	4	13.8	493	1	YNN4_YEAST	HYPOTHETICAL 54.9 KD G	2.22e+03	785	4	13.8	594	1	YHYA_BPH44	HYPOTHETICAL 65 KD PRO	2.22e+03
713	4	13.8	493	1	YH2_RHOSO	HYPOTHETICAL 53.0 KD G	2.22e+03	786	4	13.8	594	1	YNN8_CAEEL	HYPOTHETICAL 66.9 KD P	2.22e+03
714	4	13.8	493	1	YEL_VICSA	VACUOLAR PROCESSING EN	2.22e+03	787	4	13.8	595	1	YV1_BPCHP	PROTEIN VP1 (OREF1)	2.22e+03
715	4	13.8	496	1	YFHK_ECOLI	PROBABLE SENSOR PROTEI	2.22e+03	788	4	13.8	597	1	Y298_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
716	4	13.8	496	1	Y018_MYCGE	HYPOTHETICAL HELICASE	2.22e+03	789	4	13.8	600	1	Y719_METJA	HYPOTHETICAL ABC TRANS	2.22e+03
717	4	13.8	497	1	YF25_CAEEL	HYPOTHETICAL 56.2 KD P	2.22e+03	790	4	13.8	600	1	YG48_YEAST	HYPOTHETICAL 68.3 KD P	2.22e+03
718	4	13.8	497	1	Y929_YEAST	HYPOTHETICAL 57.1 KD P	2.22e+03	791	4	13.8	601	1	YAN4_SCHPO	HYPOTHETICAL 69.4 KD P	2.22e+03
719	4	13.8	497	1	YFC7_YEAST	HYPOTHETICAL 57.6 KD P	2.22e+03	792	4	13.8	604	1	YFIC_BACSU	HYPOTHETICAL ABC TRANS	2.22e+03
720	4	13.8	498	1	YV13_MYCTU	HYPOTHETICAL 54.4 KD P	2.22e+03	793	4	13.8	605	1	YNN8_YEAST	HYPOTHETICAL 67.5 KD P	2.22e+03
721	4	13.8	499	1	YHNO_YEAST	HYPOTHETICAL 56.5 KD P	2.22e+03	794	4	13.8	608	1	YXNC_FIBSU	ENDO-1,4-BETA-XYLANASE	2.22e+03
722	4	13.8	500	1	YDAK_YEAST	HYPOTHETICAL 56.3 KD P	2.22e+03	795	4	13.8	612	1	Y081_MYCTU	HYPOTHETICAL ABC TRANS	2.22e+03
723	4	13.8	501	1	YW05_YEAST	HYPOTHETICAL 58.0 KD P	2.22e+03	796	4	13.8	614	1	Y383_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
724	4	13.8	503	1	Y226_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03	797	4	13.8	620	1	Y886_METJA	PUTATIVE MOLYBDOPROTEIN	2.22e+03
725	4	13.8	503	1	YBHE_ECOLI	HYPOTHETICAL ABC TRANS	2.22e+03	798	4	13.8	620	1	Y241_MYCGE	HYPOTHETICAL PROTEIN M	2.22e+03
726	4	13.8	503	1	YPTJ_CAEEL	HYPOTHETICAL 58.3 KD P	2.22e+03	799	4	13.8	621	1	YV40_HSVBC	CAPSID PROTEIN P40 [CO	2.22e+03
727	4	13.8	504	1	YM68_YEAST	HYPOTHETICAL 56.2 KD P	2.22e+03	800	4	13.8	623	1	Y014_MYCGE	HYPOTHETICAL ABC TRANS	2.22e+03
728	4	13.8	504	1	Y795_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	801	4	13.8	623	1	Y071_MYCTU	HYPOTHETICAL 67.7 KD P	2.22e+03
729	4	13.8	505	1	Y11R_YEAST	HYPOTHETICAL 58.3 KD P	2.22e+03	802	4	13.8	623	1	Y014_MYCPN	HYPOTHETICAL 71.4 KD P	2.22e+03
730	4	13.8	507	1	YS15_TRYBB	VARIANT SURFACE GLYCOP	2.22e+03	803	4	13.8	624	1	YMO8_YEAST	HYPOTHETICAL 71.4 KD P	2.22e+03
731	4	13.8	508	1	YHNE_ECOLI	HYPOTHETICAL 56.3 KD P	2.22e+03	804	4	13.8	630	1	YHX1_YEAST	HYPOTHETICAL 71.4 KD P	2.22e+03
732	4	13.8	512	1	Y00K_MYCTU	HYPOTHETICAL 53.6 KD P	2.22e+03	805	4	13.8	633	1	YXCC_HUMAN	DNA-REPAIR PROTEIN XRC	2.22e+03
733	4	13.8	512	1	YF56_YEAST	HYPOTHETICAL 56.6 KD P	2.22e+03	806	4	13.8	635	1	YGC3_YEAST	HYPOTHETICAL 70.6 KD P	2.22e+03
734	4	13.8	513	1	Y017_MYCTU	HYPOTHETICAL 56.3 KD P	2.22e+03	807	4	13.8	636	1	YNN6_YEAST	HYPOTHETICAL 67.4 KD P	2.22e+03
735	4	13.8	515	1	YCW2_YEAST	HYPOTHETICAL 57.0 KD T	2.22e+03	808	4	13.8	643	1	YV40_HSV2	CAPSID PROTEIN P40 [CO	2.22e+03
736	4	13.8	516	1	YVNF_AZOCH	HYPOTHETICAL PROTEIN I	2.22e+03	809	4	13.8	644	1	YV74_NPVOP	P74 PROTEIN.	2.22e+03
737	4	13.8	516	1	Y067_MYCGE	HYPOTHETICAL LIPOPROTE	2.22e+03	810	4	13.8	645	1	YV74_NPVAC	P74 PROTEIN.	2.22e+03
738	4	13.8	517	1	YXLB_BUTFI	BETA-XYLOSIDASE [EC 3.	2.22e+03	811	4	13.8	645	1	YV74_NPVCF	P74 PROTEIN.	2.22e+03
739	4	13.8	519	1	YRV8_CAEEL	PUTATIVE CYTOCHROME P4	2.22e+03	812	4	13.8	645	1	Y081_CAEEL	HYPOTHETICAL GTP-BINDI	2.22e+03
740	4	13.8	520	1	YB77_YEAST	HYPOTHETICAL 57.9 KD P	2.22e+03	813	4	13.8	645	1	Y002_MYCTU	HYPOTHETICAL 69.2 KD P	2.22e+03
741	4	13.8	524	1	YM64_YEAST	HYPOTHETICAL 56.8 KD P	2.22e+03	814	4	13.8	653	1	Y07S_MYCTU	HYPOTHETICAL 69.5 KD P	2.22e+03
742	4	13.8	524	1	Y395_MYCGE	HYPOTHETICAL LIPOPROTE	2.22e+03	815	4	13.8	655	1	YAGF_ECOLI	HYPOTHETICAL 69.4 KD P	2.22e+03
743	4	13.8	525	1	YBFA_YEAST	HYPOTHETICAL 59.2 KD P	2.22e+03	816	4	13.8	655	1	YHCP_ECOLI	HYPOTHETICAL 73.6 KD P	2.22e+03
744	4	13.8	527	1	YEGH_ECOLI	HYPOTHETICAL 59.5 KD P	2.22e+03	817	4	13.8	655	1	YKDA_MYCCA	HYPOTHETICAL 75.9 KD P	2.22e+03
745	4	13.8	529	1	YSN6_TRYBB	VARIANT SURFACE GLYCOP	2.22e+03	818	4	13.8	661	1	Y182_SYNV3	HYPOTHETICAL ABC TRANS	2.22e+03
746	4	13.8	530	1	YK03_YEAST	HYPOTHETICAL 60.8 KD P	2.22e+03	819	4	13.8	665	1	YXBO_YEAST	HYPOTHETICAL 76.8 KD P	2.22e+03
747	4	13.8	533	1	Y057_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03	820	4	13.8	666	1	YXBO_YEAST	HYPOTHETICAL 77.7 KD P	2.22e+03
748	4	13.8	535	1	YNNB_BACPU	BETA-XYLOSIDASE [EC 3.	2.22e+03	821	4	13.8	667	1	Y366_MYCGE	PROBABLE TONB-DEPENDEN	2.22e+03
749	4	13.8	536	1	Y0U1_CAEEL	HYPOTHETICAL 59.1 KD P	2.22e+03	822	4	13.8	670	1	YF67_HAEIN	PROBABLE TONB-DEPENDEN	2.22e+03
750	4	13.8	536	1	YRN3_CAEEL	HYPOTHETICAL 61.0 KD P	2.22e+03	823	4	13.8	670	1	YX8_SCHPO	HYPOTHETICAL 75.2 KD P	2.22e+03
751	4	13.8	542	1	YIO6_YEAST	PUTATIVE TRANSPORTER Y	2.22e+03	824	4	13.8	674	1	YB66_YEAST	HYPOTHETICAL 77.7 KD P	2.22e+03
752	4	13.8	543	1	YDE3_SCHPO	HYPOTHETICAL 61.8 KD P	2.22e+03	825	4	13.8	676	1	Y043_YEAST	HYPOTHETICAL 77.0 KD P	2.22e+03
753	4	13.8	544	1	Y524_CAEEL	PUTATIVE CYTOCHROME P4	2.22e+03	826	4	13.8	676	1	YNEH_CAEEL	YME1 PROTEIN HOMOLOG (	2.22e+03

827	4	13.8	684	1	YBC4_YEAST	HYPOTHETICAL 77.9 KD P	2.22e+03	900	4	13.8	885	1	Y143_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
828	4	13.8	685	1	YPR4_CAEEL	HYPOTHETICAL 77.0 KD T	2.22e+03	901	4	13.8	889	1	Y188_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
829	4	13.8	687	1	YLI18_CAEEL	HYPOTHETICAL 76.7 KD P	2.22e+03	902	4	13.8	908	1	Y27_BRD	STRUCTURAL CORE PROTI	2.22e+03
830	4	13.8	687	1	YBG0_YEAST	HYPOTHETICAL 78.8 KD P	2.22e+03	903	4	13.8	910	1	Y068_CAEEL	HYPOTHETICAL 105.6 KD	2.22e+03
831	4	13.8	688	1	YEO7_YEAST	HYPOTHETICAL 79.5 KD P	2.22e+03	904	4	13.8	913	1	Y20_CAEEL	HYPOTHETICAL 104.9 KD	2.22e+03
832	4	13.8	692	1	YQIR_BACSU	PUTATIVE SIGMA L-DEPEN	2.22e+03	905	4	13.8	921	1	Y596_SCHPO	HYPOTHETICAL 104.4 KD	2.22e+03
833	4	13.8	697	1	Y0F8_MCTU	HYPOTHETICAL ABC TRANS	2.22e+03	906	4	13.8	923	1	YEL4_YEAST	HYPOTHETICAL 103.3 KD	2.22e+03
834	4	13.8	704	1	YPS1_YEAST	VACUOLAR SORTING PROTE	2.22e+03	907	4	13.8	943	1	YMI1_YEAST	HYPOTHETICAL 105.9 KD	2.22e+03
835	4	13.8	706	1	YEU3_YEAST	HYPOTHETICAL 81.5 KD P	2.22e+03	908	4	13.8	958	1	Y478_SCHPO	HYPOTHETICAL 107.1 KD	2.22e+03
836	4	13.8	719	1	YMP8_CAEEL	HYPOTHETICAL 82.6 KD P	2.22e+03	909	4	13.8	962	1	YAXA_SCHPO	HYPOTHETICAL 112.7 KD	2.22e+03
837	4	13.8	719	1	YK44_CAEEL	HYPOTHETICAL 82.8 KD P	2.22e+03	910	4	13.8	968	1	YV34_MCTU	HYPOTHETICAL 106.2 KD	2.22e+03
838	4	13.8	721	1	YI31_YEAST	HYPOTHETICAL 84.0 KD P	2.22e+03	911	4	13.8	968	1	Y682_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
839	4	13.8	723	1	Y262_HAEIN	PROBABLE TONB-DEPENDEN	2.22e+03	912	4	13.8	971	1	VP2_EDV1	OUTER CAPSID PROTEIN V	2.22e+03
840	4	13.8	726	1	YCC6_ECOLI	HYPOTHETICAL 81.2 KD P	2.22e+03	913	4	13.8	978	1	YX7_CAEEL	HYPOTHETICAL 113.1 KD	2.22e+03
841	4	13.8	728	1	YB7_YEAST	HYPOTHETICAL 83.0 KD P	2.22e+03	914	4	13.8	983	1	Y144_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
842	4	13.8	732	1	YJM1_CAEEL	HYPOTHETICAL 81.8 KD P	2.22e+03	915	4	13.8	989	1	YD30_YEAST	HYPOTHETICAL 112.2 KD	2.22e+03
843	4	13.8	732	1	YAGR_ECOLI	HYPOTHETICAL 78.1 KD P	2.22e+03	916	4	13.8	990	1	YAA2_SCHPO	HYPOTHETICAL 111.5 KD	2.22e+03
844	4	13.8	735	1	YKH8_YEAST	PROBABLE ATP-DEPENDENT	2.22e+03	917	4	13.8	993	1	YAJ1_SCHPO	PUTATIVE FAMILY 31 GLU	2.22e+03
845	4	13.8	736	1	YBL7_YEAST	HYPOTHETICAL 82.4 KD P	2.22e+03	918	4	13.8	995	1	YD1_CAEEL	HYPOTHETICAL 111.2 KD	2.22e+03
846	4	13.8	738	1	YAS9_SCHPO	HYPOTHETICAL 82.4 KD P	2.22e+03	919	4	13.8	995	1	YD31_YEAST	PUTATIVE ATP-DEPENDENT	2.22e+03
847	4	13.8	738	1	YK74_YEAST	HYPOTHETICAL 84.0 KD P	2.22e+03	920	4	13.8	997	1	YX2_CAEEL	HYPOTHETICAL 110.3 KD	2.22e+03
848	4	13.8	738	1	Y013_BPL2	HYPOTHETICAL 81.3 KD P	2.22e+03	921	4	13.8	1011	1	VP3_YEAST	VACUOLAR PROTEIN SORTI	2.22e+03
849	4	13.8	742	1	ZFA_MOUSE	ZINC FINGER AUTOSOMAL	2.22e+03	922	4	13.8	1019	1	VP3_RDV	MAJOR 114 KD STRUCTURA	2.22e+03
850	4	13.8	748	1	Y473_SNY3	HYPOTHETICAL 84.2 KD P	2.22e+03	923	4	13.8	1026	1	YH27_CAEEL	PUTATIVE HELICASE W08D	2.22e+03
851	4	13.8	758	1	YJ06_YEAST	PUTATIVE 86.7 KD TRANS	2.22e+03	924	4	13.8	1029	1	YFC5_YEAST	HYPOTHETICAL 117.8 KM	2.22e+03
852	4	13.8	758	1	YM38_YEAST	HYPOTHETICAL 85.0 KD P	2.22e+03	925	4	13.8	1030	1	Y075_MYCPN	HYPOTHETICAL PROTEIN M	2.22e+03
853	4	13.8	758	1	YLDM_CAEEL	HYPOTHETICAL 85.0 KD P	2.22e+03	926	4	13.8	1036	1	Y414_MYCGE	HYPOTHETICAL XYLANASE	2.22e+03
854	4	13.8	759	1	YCD9_YEAST	HYPOTHETICAL 86.0 KD P	2.22e+03	927	4	13.8	1035	1	YXNA_THENE	ENDO-1,4-BETA-XYLANASE	2.22e+03
855	4	13.8	760	1	YCD5_YEAST	HYPOTHETICAL 86.6 KD P	2.22e+03	928	4	13.8	1059	1	YXNA_THEMA	ENDO-1,4-BETA-XYLANASE	2.22e+03
856	4	13.8	760	1	YAC1_YEAST	HYPOTHETICAL 87.2 KD P	2.22e+03	929	4	13.8	1060	1	YNI8_YEAST	HYPOTHETICAL 118.3 KD	2.22e+03
857	4	13.8	765	1	Y042_MCTU	HYPOTHETICAL 87.7 KD P	2.22e+03	930	4	13.8	1073	1	Y126_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
858	4	13.8	767	1	YNY7_YEAST	PUTATIVE CYSTEINYL-TRN	2.22e+03	931	4	13.8	1075	1	Y124_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
859	4	13.8	767	1	YMW4_YEAST	HYPOTHETICAL 86.9 KD P	2.22e+03	932	4	13.8	1077	1	YXNY_CLOTG	ENDO-1,4-BETA-XYLANASE	2.22e+03
860	4	13.8	772	1	Y404_SCHPO	HYPOTHETICAL 86.2 KD P	2.22e+03	933	4	13.8	1084	1	Y172_HAEIN	PROBABLE TONB-DEPENDEN	2.22e+03
861	4	13.8	776	1	VP4_ROPB4	OUTER CAPSID PROTEIN V	2.22e+03	934	4	13.8	1085	1	YAF4_SCHPO	HYPOTHETICAL 122.9 KD	2.22e+03
862	4	13.8	778	1	YF88_YEAST	HYPOTHETICAL 88.7 KD H	2.22e+03	935	4	13.8	1087	1	YXNX_CLOTG	EXOGUCANASE XYNX PREC	2.22e+03
863	4	13.8	780	1	YAB3_SCHPO	HYPOTHETICAL 89.2 KD H	2.22e+03	936	4	13.8	1095	1	YOH5_YEAST	PROBABLE ATP-DEPENDENT	2.22e+03
864	4	13.8	782	1	YIN2_YEAST	HYPOTHETICAL 88.8 KD P	2.22e+03	937	4	13.8	1105	1	YE6E_ECOLI	HYPOTHETICAL 123.9 KD	2.22e+03
865	4	13.8	782	1	YMG2_YEAST	HYPOTHETICAL 89.2 KD P	2.22e+03	938	4	13.8	1113	1	Y140_MYCPN	HYPOTHETICAL ATP-BINDI	2.22e+03
866	4	13.8	786	1	YD59_SCHPO	HYPOTHETICAL PROTEIN C	2.22e+03	939	4	13.8	1113	1	Y140_MYCGE	HYPOTHETICAL ATP-BINDI	2.22e+03
867	4	13.8	786	1	YM01_CAEEL	HYPOTHETICAL 88.1 KD P	2.22e+03	940	4	13.8	1118	1	YIP1_YEAST	HYPOTHETICAL 128.7 KD	2.22e+03
868	4	13.8	788	1	YG4C_YEAST	HYPOTHETICAL 89.4 KD T	2.22e+03	941	4	13.8	1129	1	YB95_YEAST	HYPOTHETICAL 131.1 KD	2.22e+03
869	4	13.8	794	1	YTB7_BTF7	TAIL TUBULAR PROTEIN B	2.22e+03	942	4	13.8	1139	1	YMB3_CAEEL	PROBABLE INTEGRIN ALPH	2.22e+03
870	4	13.8	796	1	YAB9_SCHPO	HYPOTHETICAL 90.9 KD P	2.22e+03	943	4	13.8	1157	1	YXNA_THESA	ENDO-1,4-BETA-XYLANASE	2.22e+03
871	4	13.8	796	1	YRS7_CAEEL	HYPOTHETICAL 90.8 KD P	2.22e+03	944	4	13.8	1166	1	YMRK_XIPMA	MELANOMA RECEPTOR PROT	2.22e+03
872	4	13.8	805	1	YPH1_SCHPO	PROBABLE VACUOLAR ATP	2.22e+03	945	4	13.8	1193	1	YIPR_BACSU	HYPOTHETICAL 137.4 KD	2.22e+03
873	4	13.8	807	1	YMT9_YEAST	HYPOTHETICAL 91.7 KD T	2.22e+03	946	4	13.8	1196	1	XPG_XENLA	DNA-REPAIR PROTEIN COM	2.22e+03
874	4	13.8	807	1	YB9T_YEAST	PROBABLE PROTEIN-TYROS	2.22e+03	947	4	13.8	1202	1	YE01_SCHPO	HYPOTHETICAL 138.5 KD	2.22e+03
875	4	13.8	808	1	Y066_NPVAC	HYPOTHETICAL 94.0 KD P	2.22e+03	948	4	13.8	1203	1	XCP2_XENLA	CHROMOSOME ASSEMBLY PR	2.22e+03
876	4	13.8	819	1	YC81_CHLVO	HYPOTHETICAL 94.7 KD P	2.22e+03	949	4	13.8	1213	1	YXN5_CAEEL	HYPOTHETICAL 139.4 KD	2.22e+03
877	4	13.8	821	1	YQIG_ECOLI	HYPOTHETICAL 94.7 KD P	2.22e+03	950	4	13.8	1225	1	Y309_MYCGE	HYPOTHETICAL LIPOPROTE	2.22e+03
878	4	13.8	827	1	Y4L4_RHSN	HYPOTHETICAL 91.8 KD K	2.22e+03	951	4	13.8	1237	1	YG2L_YEAST	HYPOTHETICAL 140.5 KD	2.22e+03
879	4	13.8	830	1	YMG2_YEAST	HYPOTHETICAL 94.9 KD P	2.22e+03	952	4	13.8	1237	1	YD2_SCHPO	HYPOTHETICAL 143.0 KD	2.22e+03
880	4	13.8	831	1	YMG3_CAEEL	HYPOTHETICAL 95.8 KD P	2.22e+03	953	4	13.8	1244	1	Y307_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03
881	4	13.8	831	1	YVPI_HUMAN	CLATHRIN-COATED VESICL	2.22e+03	954	4	13.8	1250	1	YFAL_ECOLI	HYPOTHETICAL 131.2 KD	2.22e+03
882	4	13.8	838	1	VPPI_RAT	CLATHRIN-COATED VESICL	2.22e+03	955	4	13.8	1272	1	Y228_METJA	HYPOTHETICAL PROTEIN M	2.22e+03
883	4	13.8	838	1	VPPI_BOVIN	HYPOTHETICAL 97.0 KD P	2.22e+03	956	4	13.8	1273	1	WEB1_YEAST	WEB1 PROTEIN (PROTEIN	2.22e+03
884	4	13.8	839	1	YNE1_YEAST	HYPOTHETICAL 95.2 KD P	2.22e+03	957	4	13.8	1277	1	YMW6_YEAST	HYPOTHETICAL 147.0 KD	2.22e+03
885	4	13.8	839	1	YR86_CAEEL	HYPOTHETICAL 95.2 KD P	2.22e+03	958	4	13.8	1293	1	XPC_DROME	DNA-REPAIR PROTEIN COM	2.22e+03
886	4	13.8	848	1	YB88_YEAST	HYPOTHETICAL 92.8 KD P	2.22e+03	959	4	13.8	1299	1	VP75_HSVSA	PROBABLE MEMBRANE ANTI	2.22e+03
887	4	13.8	849	1	Y343_METJA	HYPOTHETICAL PROTEIN M	2.22e+03	960	4	13.8	1300	1	Y338_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03
888	4	13.8	852	1	YKMI_YEAST	HYPOTHETICAL 96.0 KD T	2.22e+03	961	4	13.8	1318	1	YDH4_SCHPO	HYPOTHETICAL 150.9 KD	2.22e+03
889	4	13.8	854	1	VP41_MYCTU	VACUOLAR ASSEMBLY PROT	2.22e+03	962	4	13.8	1322	1	YAG3_YEAST	HYPOTHETICAL 138.1 KD	2.22e+03
890	4	13.8	859	1	Y07C_HUMAN	HYPOTHETICAL ABC TRANS	2.22e+03	963	4	13.8	1325	1	Y309_MYCPN	HYPOTHETICAL LIPOPROTE	2.22e+03
891	4	13.8	863	1	YS97_CAEEL	HYPOTHETICAL 97.2 KD P	2.22e+03	964	4	13.8	1332	1	XKDO_BACSU	PHASE-LIKE ELEMENT PBS	2.22e+03
892	4	13.8	870	1	YCS0_YEAST	HYPOTHETICAL 96.1 KD P	2.22e+03	965	4	13.8	1351	1	YOL4_CAEEL	HYPOTHETICAL 152.4 KD	2.22e+03
893	4	13.8	871	1	Y47D_SCHPO	HYPOTHETICAL 98.4 KD P	2.22e+03	966	4	13.8	1357	1	YJ03_YEAST	HYPOTHETICAL 154.9 KD	2.22e+03
894	4	13.8	875	1	VP34_YEAST	PHOSPHATIDYLINOSITOL 3	2.22e+03	967	4	13.8	1401	1	WRN_MOUSE	WERNER SYNDROME HELICA	2.22e+03
895	4	13.8	875	1	Y066_NPVOP	HYPOTHETICAL 98.6 KD P	2.22e+03	968	4	13.8	1403	1	YGN1_YEAST	HYPOTHETICAL 163.2 KD	2.22e+03
896	4	13.8	881	1	YFCU_ECOLI	HYPOTHETICAL OUTER MEM	2.22e+03	969	4	13.8	1403	1	YDF3_SCHPO	PROBABLE EUKARYOTIC IN	2.22e+03
897	4	13.8	881	1	Y310_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03	970	4	13.8	1478	1	YAWB_SCHPO	PROBABLE ATP-DEPENDENT	2.22e+03
898	4	13.8	881	1	YJ78_YEAST	HYPOTHETICAL 97.7 KD M	2.22e+03	971	4	13.8	1507	1	Y056_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03
899	4	13.8	882	1	YNK1_CAEEL	HYPOTHETICAL 98.3 KD P	2.22e+03	972	4	13.8	1524	1	Y133_HUMAN	HYPOTHETICAL PROTEIN K	2.22e+03

973 4 13.8 1549 1 YW2\_YEAST PUTATIVE MEMBRANE GLYC 2.22e+03  
974 4 13.8 1549 1 YR3\_YEAST PUTATIVE MEMBRANE GLYC 2.22e+03  
975 4 13.8 1551 1 Y012\_CAEEL HYPOTHETICAL 175.7 KD 2.22e+03  
976 4 13.8 1559 1 Y115\_YEAST PROBABLE ATP-DEPENDENT 2.22e+03  
977 4 13.8 1560 1 X169\_HUMAN XE169 PROTEIN (SMCX PR 2.22e+03  
978 4 13.8 1569 1 YFJA\_ECOLI HYPOTHETICAL 98.4 KD P 2.22e+03  
979 4 13.8 1658 1 YM67\_YEAST HYPOTHETICAL 187.1 KD 2.22e+03  
980 4 13.8 1661 1 YBT1\_YEAST ATP-DEPENDENT BILE ACI 2.22e+03  
981 4 13.8 1687 1 YDHA\_SCHPO HYPOTHETICAL 192.5 KD 2.22e+03  
982 4 13.8 1693 1 Y163\_SYNY3 HYPOTHETICAL WD-REPEAT 2.22e+03  
983 4 13.8 1698 1 Y076\_HUMAN HYPOTHETICAL PROTEIN K 2.22e+03  
984 4 13.8 1745 1 Z01\_MOUSE TIGHT JUNCTION PROTEIN 2.22e+03  
985 4 13.8 1788 1 YF72\_CAEEL HYPOTHETICAL 200.6 KD 2.22e+03  
986 4 13.8 1799 1 Y025\_CAEEL HYPOTHETICAL 202.6 KD 2.22e+03  
987 4 13.8 1807 1 YV2\_XENLA VITELLOGENIN A2 PRECUR 2.22e+03  
988 4 13.8 1882 1 Y468\_MICPN HYPOTHETICAL PROTEIN M 2.22e+03  
989 4 13.8 1942 1 Y054\_HUMAN HYPOTHETICAL PROTEIN K 2.22e+03  
990 4 13.8 2131 1 YCF2\_SPIOI HYPOTHETICAL 250 KD PR 2.22e+03  
991 4 13.8 2136 1 YCF2\_MARPO HYPOTHETICAL 259 KD PR 2.22e+03  
992 4 13.8 2216 1 YCF2\_EPIVI HYPOTHETICAL 260 KD PR 2.22e+03  
993 4 13.8 2334 1 WAPA\_BACSU WALL-ASSOCIATED PROTEI 2.22e+03  
994 4 13.8 2493 1 YBA4\_YEAST HYPOTHETICAL 287.5 KD 2.22e+03  
995 4 13.8 2688 1 ZEP1\_MOUSE ZINC FINGER PROTEIN 40 2.22e+03  
996 4 13.8 2717 1 ZEP1\_HUMAN ZINC FINGER PROTEIN 40 2.22e+03  
997 4 13.8 2813 1 VWF\_HUMAN VON WILLEBRAND FACTOR 2.22e+03  
998 4 13.8 3051 1 YXK3\_CAEEL HYPOTHETICAL PROTEIN T 2.22e+03  
999 4 13.8 3144 1 VP13\_YEAST VACUOLAR PROTEIN SORTI 2.22e+03  
1000 4 13.8 4385 1 YP73\_CAEEL HYPOTHETICAL 486.0 KD 2.22e+03

## ALIGNMENTS

RESULT 1 STANDARD: PRT: 999 AA.  
ID DSG3\_HUMAN  
AC P32926;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE DESMOGLEIN 3 PRECURSOR (130 KD PEMPHIGUS VULGARIS ANTIGEN) (PVA).  
GN DSG3.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RX MEDLINE; 92069753.  
RA ANAGAI M., KLAUS-KOVTON V., STANLEY J.R.;  
RT "Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a disease of cell adhesion."  
RL CELL 67:869-877(1991).  
CC -!- FUNCTION: COMPONENT OF INTERCELLULAR DESMOSOME JUNCTIONS.  
CC INVOLVED IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE  
CC FILAMENTS MEDIATING CELL-CELL ADHESION.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: EPIDERMIS, TONGUE, TONSIL, OESOPHAGUS AND  
CC CARCINOMAS.  
CC -!- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS  
CC (POTENTIAL).  
CC -!- DISEASE: PEMPHIGUS VULGARIS (PV) IS A POTENTIALLY LETHAL SKIN  
CC DISEASE IN WHICH EPIDERMAL BLISTERS OCCUR AS THE RESULT OF THE  
CC LOSS OF CELL-CELL ADHESION CAUSED BY THE ACTION OF AUTOANTIBODIES  
CC AGAINST DSG3.  
CC -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY. BELONGS TO THE  
CC DESMOSOMAL SUBFAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M76482; G190752; -  
DR PIR; A41088; IJHUG3.  
DR MIM; 169615; -  
DR PROSITE; PS00232; CADHERIN; 3.  
DR PFAM; PF00028; cadherin; 4.  
DR HSSP; P09803; 1EDH.  
KW CELL ADHESION; SIGNAL; TRANSMEMBRANE; CYTOSKELETON; GLYCOPROTEIN;  
KW CALCIUM-BINDING; REPEAT.  
FT SIGNAL 1 23 POTENTIAL.  
FT PROPEP 24 49 POTENTIAL.  
FT CHAIN 50 999 DESMOGLEIN 3.  
FT DOMAIN 50 615 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 616 640 POTENTIAL.  
FT DOMAIN 641 999 CYTOPLASMIC (POTENTIAL).  
FT REPEAT 50 158 CADHERIN 1.  
FT REPEAT 159 268 CADHERIN 2.  
FT REPEAT 269 383 CADHERIN 3.  
FT REPEAT 386 499 CADHERIN 4.  
FT REPEAT 910 935 DESMOGLEIN REPEAT 1.  
FT REPEAT 936 966 DESMOGLEIN REPEAT 2.  
FT CARBOHYD 110 110 POTENTIAL.  
FT CARBOHYD 180 180 POTENTIAL.  
FT CARBOHYD 459 459 POTENTIAL.  
FT CARBOHYD 545 545 POTENTIAL.  
SQ SEQUENCE 999 AA; 107503 MW; 4891F6AE CRC32;

Query Match 24.18; Score 7; DB 1; Length 999;

Best Local Similarity 100.0%; Pred. No. 2.08e-02; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

Db 100 FVVDKNT 106  
||||||

QY 22 FVVDKNT 28

RESULT 2  
ID VATF\_MANSE STANDARD: PRT: 124 AA.  
AC P31478;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE VACUOLAR ATP SYNTHASE 14 KD SUBUNIT (EC 3.6.1.34) (V-ATPASE F  
DE SUBUNIT).  
GN VHA14.  
OS MANDUCA SEXTA (TOBACCO HAWKMOH) (TOBACCO HORNWORM).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; LEPIDOPTERA; SPHINGIODEA; SPHINGINAE; MANDUCA.  
RN [1]  
RX MEDLINE; 9414885.  
RA GRAEF R., LEPIER A., HARVEY W.R., WIECZOREK H.;  
RT "A novel 14-kDa V-ATPase subunit in the tobacco hornworm midgut."  
RL J. BIOL. CHEM. 269:3767-3774(1994).  
CC -!- FUNCTION: THE VACUOLAR ATPASE IS RESPONSIBLE FOR ACIDIFYING A  
CC VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC CELLS. V-  
CC ATPASE IS RESPONSIBLE FOR ENERGIZING ELECTROPHORETIC K+/2H+  
CC ANTIPORT BY GENERATING A TRANSMEMBRANE VOLTAGE OF MORE THAN 200MV.  
CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A  
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,  
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE  
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).  
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS V-ATPASE SUBUNIT F,  
CC V-TYPE SODIUM ATPASE NTFG AND ARCHEAL ATPASE SUBUNIT F.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DR EMBL: X67130; G9729; -.
DR PIR: S25013; S25013.
KW HYDROLASE; HYDROGEN ION TRANSPORT.
SQ SEQUENCE 124 AA; 13814 MW; 4C4A5560 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 VVDKNT 47
QY 23 VVDKNT 28

RESULT 3
ID VATE DROME STANDARD; PRT: 124 AA.
AC Q24583;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE VACUOLAR ATP SYNTHASE 14 KD SUBUNIT (EC 3.6.1.34) (V-ATPASE F
DE SUBUNIT).
GN VH414.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEAD.
RX MEDLINE: 96269411.
RA GUO Y., KAISER K., WIECZOREK H., DOW J.A.T.;
RT "The Drosophila melanogaster gene vha14 encoding a 14-kDa F-subunit
RL GENE 172:239-243(1996).
CC -!- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE
CC ESSENTIAL FOR ASSEMBLY OR CATALYTIC FUNCTION. V-ATPASE IS
CC RESPONSIBLE FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS
CC IN EUKARYOTIC CELLS.
CC -!- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNIT A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN).
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS V-ATPASE SUBUNIT F,
CC V-TYPE SODIUM ATPASE NTPG AND ARCHEAL ATPASE SUBUNIT F.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69379; E226035; -.
KW HYDROLASE; HYDROGEN ION TRANSPORT.
SQ SEQUENCE 127 AA; 14162 MW; E0DA416F CRC32;

Query Match 20.7%; Score 6; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 VVDKNT 47
QY 23 VVDKNT 28

RESULT 5
ID SUCA HUMAN STANDARD; PRT: 132 AA.
AC P53597;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE SUCCINYL-COA LIGASE [GDP-FORMING], ALPHA-CHAIN PRECURSOR (EC 6.2.1.4)
DE (SUCCINYL-COA SYNTHETASE, ALPHA CHAIN) (SCS-ALPHA) (FRAGMENT).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND SKELETAL MUSCLE.
RA JAMES M., MAN N., EDWARDS Y.H., MORRIS G.E.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: SUCCINATE + COA + GTP -> SUCCINYL-COA + GDP +
CC ORTHOPHOSPHATE.
CC -!- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC ACID CYCLE.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- SIMILARITY: TO OTHER ALPHA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; Z68204; E213912; -  
DR PROSITE; PS01216; SUCCINYL\_COA\_LIG\_1; PARTIAL.  
DR PROSITE; PS00399; SUCCINYL\_COA\_LIG\_2; PARTIAL.  
DR HSP9; P07459; 1SCU.  
KW LIGASE; GLYCOLYSIS; TRICARBOXYLIC ACID CYCLE; MITOCHONDRION;  
KW TRANSIT PEPTIDE; GTP-BINDING; PHOSPHORYLATION.  
FT TRANSIT 1 27 MITOCHONDRION (BY SIMILARITY).  
FT CHAIN 28 >132 SUCCINYL-COA LIGASE [GDP-FORMING], ALPHA-  
FT CHAIN.

FT NON\_TER 132 132  
SQ SEQUENCE 132 AA; 13961 MW; F056EE32 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1.67e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 VDKNTK 44  
|||||  
Qy 24 VDKNTK 29

RESULT 6  
ID AR16\_YEAST STANDARD; PRT; 154 AA.  
AC P40518;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE ARP2/3 COMPLEX 16 KD SUBUNIT (P16-ARC).  
GN ARC15 OR YIL062C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA BARRELL B.G., BADCOCK K., BANKIER A.T., BOWMAN S., BROWN D.,  
RA CHURCHER C.M., COPSEY T., DEAR S., DEVLIN K., FRASER A.,  
RA GENTLES S., HAMLIN N., HORSNELL T.S., HUNT S., JAGELS K., JONES M.,  
RA LOUIS E., LYE G., MOULE S., MOULE T., ODELL C., PEARSON D.,  
RA RAJANDREAM M.A., RILES L., ROWLEY N., SKELTON J., SMITH V.,  
RA WALSH S.V., WHITEHEAD S.;  
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -!- FUNCTION: PART OF A COMPLEX IMPLICATED IN THE CONTROL OF ACTIN  
CC POLYMERIZATION IN CELLS (BY SIMILARITY).  
CC -!- SUBUNIT: BELONGS TO A COMPLEX COMPOSED OF ARP2, ARP3, P41-ARC,  
CC P34-ARC, P21-ARC, P20-ARC AND P16-ARC (BY SIMILARITY).  
CC -!- SIMILARITY: STRONG, TO OTHER P16-ARC SUBUNITS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; Z38060; G557807; -  
DR EMBL; Z47047; G763284; -  
DR PIR; S48417; S48417;  
DR SGD; L0004032; ARC15.  
SQ SEQUENCE 154 AA; 17134 MW; B1FFB5CD CRC32;

Query Match 20.7%; Score 6; DB 1; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.67e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 GQKGG 126  
|||||

Qy 16 GQKGG 21

RESULT 7  
ID FLRE\_BOVIN STANDARD; PRT; 205 AA.  
AC P25556;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE FLAVIN REDUCTASE (EC 1.6.99.1) (FR) (NADPH-DEPENDENT DIAPHORASE)  
DE (NADPH-FLAVIN REDUCTASE) (FLR) (BILIVERDIN REDUCTASE B) (EC 1.3.1.24)  
DE (BVR-B) (BILIVERDIN-IX BETA-REDUCTASE) (GREEN HEME BINDING PROTEIN)  
DE (GHBP).  
GN BLVRB.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; BOS.  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-57 AND 117-126.  
RC TISSUE=LIVER;  
RX MEDLINE; 95023903.  
RA QUANT K.S., HULTQUIST D.E.;  
RT "Flavin reductase: sequence of cDNA from bovine liver and tissue  
RT distribution."  
RL PROC. NATL. ACAD. SCI. U.S.A. 91:9322-9326(1994).  
RN [2]  
RP SEQUENCE OF 1-36; 120-136 AND 145-168.  
RC TISSUE=ERYTHROCYTE;  
RX MEDLINE; 91298967.

RA QUANT K.S., XU F., CHEN P., HULTQUIST D.E.;  
RT "Evidence that the protein components of bovine erythrocyte green  
RT heme binding protein and flavin reductase are identical."  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 178:315-321(1991).  
CC -!- FUNCTION: CATALYZES ELECTRON TRANSFER FROM REDUCED PYRIDINE  
CC NUCLEOTIDES TO FLAVINS AS WELL AS METHYLENE BLUE, PYRROLOQUINOLINE  
CC QUINONE, RIBOFLAVIN, OR METHEMOGLOBIN. POSSIBLE ROLE IN PROTECTING  
CC CELLS FROM OXIDATIVE DAMAGE OR IN REGULATING IRON METABOLISM. IN  
CC THE LIVER, CONVERTS BILIVERDIN TO BILIRUBIN.  
CC -!- CATALYTIC ACTIVITY: NADPH + ACCEPTOR = NADP(+) + REDUCED ACCEPTOR.  
CC -!- CATALYTIC ACTIVITY: BILIRUBIN + NAD(P)(+) = BILIVERDIN + NAD(P)H.  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
CC -!- TISSUE SPECIFICITY: AT LEAST EXPRESSED IN THE LIVER AND  
CC ERYTHROCYTE.

CC -!- MASS SPECTROMETRY: MW=21994; METHOD=ELECTROSPRAY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; L35045; G516594; -  
KW OXIDOREDUCTASE; NADP; MULTIFUNCTIONAL ENZYME.  
FT INIT\_MET 0 0  
FT NP\_BIND 4 35 NAD/NADP (POTENTIAL).  
SQ SEQUENCE 205 AA; 22001 MW; 9836F3FF CRC32;

Query Match 20.7%; Score 6; DB 1; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.67e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 113 LLWDPS 118  
|||||  
Qy 10 LLWDPS 15

RESULT 8  
ID LPXA\_RICRI STANDARD; PRT; 264 AA.  
AC P32199;  
DT 01-OCT-1993 (REL. 27, CREATED)

```
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ACYL-[ACYL-CARRIER-PROTEIN]-UDP-N-ACETYLGLUCOSAMINE O-ACYLTRANSFERASE
DE (EC 2.3.1.129) (UDP-N-ACETYLGLUCOSAMINE ACYLTRANSFERASE).
GN LPXA.
OS RICKETTSIA RICKETTSII.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RICKETTSIALES;
OC RICKETTSIAEAE; RICKETTSIAE; RICKETTSIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94171066.
RA SHAW E.I., WOOD D.O.:
RT "Characterization of a Rickettsia rickettsii DNA fragment analogous
RT to the fir A-ORF17-1pxa region of Escherichia coli.";
RL GENE 140:109-113(1994).
CC -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF LIPID A, A
CC PHOSPHORYLATED GLYCOLIPID THAT ANCHORS THE LIPOPOLYSACCHARIDE TO
CC THE OUTER MEMBRANE OF THE CELL.
CC -!- CATALYTIC ACTIVITY: (R)-3-HYDROXYTETRADECANOYL-[ACYL-CARRIER
CC PROTEIN] + UDP-N-ACETYLGLUCOSAMINE -> [ACYL-CARRIER PROTEIN] +
CC UDP-3-O-(3-HYDROXYTETRADECANOYL)-N-ACETYLGLUCOSAMINE.
CC -!- PATHWAY: FIRST STEP IN LIPID A BIOSYNTHESIS.
CC -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -!- SUBUNIT: CYTOPLASMIC.
CC -!- SIMILARITY: STRONG, TO LPXA IN OTHER BACTERIA.
CC -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NDL FAMILY OF
CC ACYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L22690; G349108; -.
CC PROSITE: PS00101; HEXAPEP_TRANSFERASES; 1.
CC PFAM: PF00132; hexapep; 4.
CC HSSP: P10440; 1LXA.
KW TRANSFERASE; ACYLTRANSFERASE; LIPID A BIOSYNTHESIS; LIPID SYNTHESIS;
KW REPEAT.
SQ SEQUENCE 264 AA; 28333 MW; 578D466 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 264;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 68 QPPQIL 73
QY 5 QPPQIL 10
|||||

RESULT 9
ID FTSZ_CAUCR STANDARD; PRT; 508 AA.
AC P52976;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN FTSZ.
GN FTSZ.
OS CAULOBACTER CRESCENTUS.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; CAULOBACTER GROUP;
OC CAULOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-NA1000;
RX MEDLINE: 96270535.
RA QUARDOKUS E., DIN N., BRUN Y.V.:
RT "Cell cycle regulation and cell type-specific localization of the
RT Ftsz division initiation protein in Caulobacter.";
RL PROC. NATL. ACAD. SCI. U.S.A. 93:6314-6319(1996).
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO THE CELL-DIVISION PROCESS.

ITS SEEMS TO ASSEMBLE INTO A DYNAMIC RING ON THE INNER SURFACE OF
THE CYTOPLASMIC MEMBRANE AT THE PLACE WHERE DIVISION WILL OCCUR,
AND THE FORMATION OF THE RING IS THE SIGNAL FOR SEPTATION TO
BEGIN. BINDS TO AND HYDROLYZES GTP (BY SIMILARITY).
-!- SUBUNIT: AGGREGATE TO FORM A RING-LIKE STRUCTURE (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSEMBLE AT THE INNER SURFACE
OF THE CYTOPLASMIC MEMBRANE (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE FTSZ FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U40273; G1304505; -.
CC PROSITE: PS01134; FTSZ_1; 1.
CC PROSITE: PS01135; FTSZ_2; 1.
CC HSSP: Q57816; 1FSZ.
KW CELL DIVISION; SEPTATION; GTP-BINDING.
FT NP_BIND 107 115 GTP (POTENTIAL).
SQ SEQUENCE 508 AA; 54221 MW; A46775F1 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 445 QKQGGF 451
QY 17 QKQGGF 22
|||||

RESULT 10
ID SDP_EIMBO STANDARD; PRT; 596 AA.
AC P42789;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE SPOZOITE DEVELOPMENTAL PROTEIN (EC 3.4.99.-).
OS EIMERIA BOVIS.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; COCCIDIA; EIMERIIDA; EIMERIIDAE;
OC EIMERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-OOCYST.
RX MEDLINE: 93149194.
RA ABRAHAMSEN M.S., CLARK T.G., MASCOLO P., SPEER C.A., WHITE M.W.:
RT "Developmental gene expression in Eimeria bovis.";
RL MOL. BIOCHEM. PARASITOL. 57:1-14(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN THE DEGRADATION OF A PROTEIN WHICH IS
CC A COMPONENT OF THE SIGNAL TRANSDUCTION PATHWAY REGULATING OOCYST
CC SPORULATION.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN SPORULATED OOCYSTS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M98842; -. NOT_ANNOTATED_CDS.
CC PROSITE: PS00143; INSULINASE; 1.
CC PFAM: PF00675; Insulinase; 1.
KW HYDROLASE; METALLOPROTEASE; ZINC.
FT METAL 77 77 ZINC (BY SIMILARITY).
FT ACT_SITE 80 80 ZINC (BY SIMILARITY).
FT METAL 81 81 ZINC (BY SIMILARITY).
```



```

FT METAL 156 156 ZINC (BY SIMILARITY).
SQ SEQUENCE 596 AA; 65279 MW; C5ECF8B3 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 596;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 336 VVDKNT 341
|||||
Qy 23 VVDKNT 28

RESULT 11
ID FIG1_MOUSE STANDARD; PRT; 630 AA.
AC O09046;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE FIG-1 PROTEIN PRECURSOR.
GN FIG1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MORIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX CHU C.C., PAUL W.E.;
RA "Fig1, an interleukin 4-induced mouse B cell gene isolated by cDNA
RT representational difference analysis.";
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2507-2512(1997).
CC -!- COFACTOR: FAD (POTENTIAL).
CC -!- INDUCTION: BY INTERLEUKIN-4.
CC -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U70429; G1924978; -.
CC DR EMBL; U70430; G1924980; -.
CC DR MGI:109552; FIG1.
CC KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; SIGNAL.
CC FT SIGNAL 1 21
CC FT CHAIN 22 630
CC FT NP_BIND 59 115
CC FT CARBOHYD 53 53
CC FT CARBOHYD 133 133
CC FT CARBOHYD 219 219
CC SQ SEQUENCE 630 AA; 70190 MW; D89F0269 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 448 QGFEV 453
|||||
Qy 19 QGFEV 24

RESULT 12
ID CAD5_PIG STANDARD; PRT; 782 AA.
AC O02840;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE VASCULAR ENDOTHELIAL-CADHERIN PRECURSOR (VE-CADHERIN) (CADHERIN-5).
GN CDH5.
OS SUS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

Query Match 20.7%; Score 6; DB 1; Length 782;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VVDKNT 124
|||||
Qy 23 VVDKNT 28

RESULT 13
ID NME4_MOUSE STANDARD; PRT; 1323 AA.
AC Q03391;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE GLUTAMATE [NMDA] RECEPTOR SUBUNIT EPSILON 4 PRECURSOR (N-METHYL
DE D-ASPARTATE RECEPTOR SUBTYPE 2D) (NR2D) (NMDAR2D).
GN GRIN2D.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
[1]
RP SEQUENCE FROM N.A.
RA KILSHAW P.J.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. THIS CADHERIN MAY PLAY A
CC IMPORTANT ROLE IN ENDOTHELIAL CELL BIOLOGY THROUGH CONTROL OF THE
CC COHESION AND ORGANIZATION OF THE INTERCELLULAR JUNCTIONS. IT
CC ASSOCIATES WITH ALPHA-CATENIN FORMING A LINK TO THE CYTOSKELETON
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. FOUND AT CELL-CELL
CC BOUNDARIES AND PROBABLY AT CELL-MATRIX BOUNDARIES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y13919; E324289; -.
CC DR PROSITE; PS00232; CADHERIN; 3.
CC DR PFAM; PF00028; cadherin; 5.
CC DR PFAM; PF01049; Cadherin_C-term; 1.
CC DR HSP; P09803; IEDH.
CC KW CELL ADHESION; GLYCOPROTEIN; PHOSPHORYLATION; TRANSMEMBRANE;
CC CALCIUM-BINDING; REPEAT; SIGNAL.
CC FT SIGNAL 1 22
CC FT PROPEP 23 44
CC FT CHAIN 45 782
CC FT DOMAIN 45 592
CC FT TRANSMEM 593 619
CC FT DOMAIN 620 782
CC FT REPEAT 45 148
CC FT REPEAT 149 255
CC FT REPEAT 256 370
CC FT REPEAT 371 475
CC FT REPEAT 476 592
CC FT DOMAIN 736 751
CC FT CARBOHYD 58 58
CC FT CARBOHYD 154 154
CC FT CARBOHYD 360 360
CC FT CARBOHYD 440 440
CC FT CARBOHYD 522 522
CC FT CARBOHYD 534 534
CC SQ SEQUENCE 782 AA; 87546 MW; 007F70E0 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 782;
Best Local Similarity 100.0%; Pred. No. 1.67e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VVDKNT 124
|||||
Qy 23 VVDKNT 28
```



OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=BRAIN;  
 RA MEDLINE; 93050214.  
 RA IKEDA K., NAGASAWA M., MORI H., ARAKI K., SAKIMURA K., WATANABE M.,  
 RA INOUE Y., MISHINA M.;  
 RT "Cloning and expression of the epsilon 4 subunit of the NMDA receptor  
 channel.";  
 RL FEBS LETT. 313:34-38(1992).  
 CC -1- FUNCTION: NMDA RECEPTOR SUBTYPE OF GLUTAMATE-GATED ION CHANNELS  
 CC POSSESSES HIGH CALCIUM PERMEABILITY AND VOLTAGE-DEPENDENT  
 CC SENSITIVITY TO MAGNESIUM AND IS MEDIATED BY GLYCINE.  
 CC SUBUNIT: HETERODIMER OF AN EPSILON SUBUNIT AND A ZETA SUBUNIT.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D12822; G220407; -  
 DR PIR; S27224; S27224.  
 DR MGD; MGI:95823; GRIN2D.  
 DR PFAM; PF00060; lig\_chan.1.  
 KW RECEPTOR; SIGNAL; TRANSMEMBRANE; POSTSYNAPTIC MEMBRANE; CALCIUM;  
 KW IONIC CHANNEL; MAGNESIUM.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 1323 GLUTAMATE [NMDA] RECEPTOR SUBUNIT  
 FT EPSILON 4.  
 FT DOMAIN 28 581 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 582 601 1 (POTENTIAL).  
 FT TRANSMEM 624 643 2 (POTENTIAL).  
 FT TRANSMEM 654 672 3 (POTENTIAL).  
 FT TRANSMEM 842 862 4 (POTENTIAL).  
 FT SITE 639 639 FUNCTIONAL DETERMINANT OF NMDA  
 FT RECEPTORS (BY SIMILARITY).  
 FT CARBOHYD 89 89 POTENTIAL.  
 FT CARBOHYD 363 363 POTENTIAL.  
 FT CARBOHYD 566 566 POTENTIAL.  
 SQ SEQUENCE 1323 AA; 143302 MW; A8B3BD2C CRC32;

Query Match 20.78; Score 6; DB 1; Length 1323;  
 Best Local Similarity 100.0%; Pred. No. 1.67e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 991 TTQPPQ 996  
 |||||  
 QY 3 TTQPPQ 8

RESULT 14  
 ID FIXU\_RHISN STANDARD; PRT; 76 AA.  
 AC Q53202;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE FIXU PROTEIN HOMOLOG.  
 GN FIXU OR Y4UJ.  
 OS RHIZOBIUM SP. (STRAIN NGR234).  
 OG PLASMID SYM.PNGR234A.  
 OC BACTERIA; PROTOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
 OC RHIZOBIACEAE; RHIZOBIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97305956.  
 RA FREIBERG C.A., FELLAY R., BAIRICH A., BROUGHTON W.J., ROSENTHAL A.,  
 RA PERRET X.;  
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";

RL NATURE 387:394-401(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96389014.  
 RA FREIBERG C., PERRET X., BROUGHTON W.J., ROSENTHAL A.;  
 RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp.  
 RT NGR234 using dye terminators and a thermostable 'sequenase': a  
 RT beginning.";  
 RL GENOME RES. 6:590-600(1996).  
 CC -1- SIMILARITY: R.LEGUMINOSARUM SYM PLASMID FIXU. ALSO TO THE  
 CC THE N-TERMINAL OF R.MELILOTI MOSB PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z68203; E213895; -  
 DR EMBL; AE000100; G2182661; -  
 KW NITROGEN FIXATION; PLASMID.  
 SQ SEQUENCE 76 AA; 8505 MW; A3FFID5E CRC32;

Query Match 17.28; Score 5; DB 1; Length 76;  
 Best Local Similarity 100.0%; Pred. No. 8.42e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 GGEVW 39  
 |||||  
 QY 20 GGEVW 24

RESULT 15  
 ID SCX2\_CENNO STANDARD; PRT; 84 AA.  
 AC P01495; Q26461;  
 DT 21-JUL-1986 (REL. 01, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE TOXIN 2 PRECURSOR (TOXIN II.9.2.2) (CN2) (FRAGMENT).  
 OS CENTRUROIDES NOXIUS (MEXICAN SCORPION).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; SCORPIONES;  
 OC BUTHOIDEA; BUTHIDAE; CENTRUROIDES.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND AMIDATION.  
 RX MEDLINE; 96119048.  
 RA VAZQUEZ A., TAPIA J.V., ELIASON W.K., MARTIN B.M., LEBRETON F.,  
 RA DELEPIERRE M., POSSANI L.D., BECERRIL B.;  
 RT "Cloning and characterization of the cDNAs encoding Na+  
 RT channel-specific toxins 1 and 2 of the scorpion Centruroides noxius  
 RT Hoffmann.";  
 RL TOXICON 33:1161-1170(1995).  
 RN [2]  
 RP SEQUENCE OF 17-82.  
 RX MEDLINE; 92155216.  
 RA ZAMUDIO F., SAAVEDRA R., MARTIN B.M., GURROLA G.B., HERION P.,  
 RA POSSANI L.D.;  
 RT "Amino acid sequence and immunological characterization with  
 RT monoclonal antibodies of two toxins from the venom of the scorpion  
 RT Centruroides noxius Hoffmann.";  
 RL EUR. J. BIOCHEM. 204:281-292(1992).  
 RN [3]  
 RP SEQUENCE OF 17-46.  
 RA MARTIN B.M., POSSANI L.D., DENT M.A.R., MAELICKE A.;  
 RT "N-terminal sequence of toxin II.9.2.2 from the venom of the Mexican  
 RT scorpion Centruroides noxius.";  
 RL TOXICON 20:74-74(1982).  
 RN [4]  
 RP DISULFIDE BONDS.  
 RX MEDLINE; 94283030.  
 RA GURROLA G.B., MORENO-HAGELSIEB G., ZAMUDSIO F.2., GARCIA M.,  
 RA SOBERON X., POSSANI L.D.;

RT "The disulfide bridges of toxin 2 from the scorpion Centruroides  
 RT noxius Hoffmann and its three-dimensional structure calculated using  
 RT the coordinates of variant 3 from Centruroides sculpturatus.";  
 RL FEBS LETT. 347:59-62(1994).  
 CC -!- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBIT THE INACTIVATION OF  
 CC THE ACTIVATED CHANNELS. THEREBY BLOCKING NEURONAL TRANSMISSION.  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.  
 CC BETA-TOXIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; S81096; G1336782; -.  
 DR PIR; A01755; NTSR2N.  
 DR PIR; S23079; S23079.  
 DR PFAM; PF00537; toxin\_3; 1.  
 DR HSP; P01494; 2SN3.  
 KW NEUROTOXIN; SODIUM CHANNEL INHIBITOR; AMIDATION; SIGNAL.  
 FT NON\_TER 1  
 FT SIGNAL <1 16  
 FT CHAIN 17 82 TOXIN 2.  
 FT DISULFID 28 81  
 FT DISULFID 32 57  
 FT DISULFID 41 62  
 FT DISULFID 45 64  
 FT MOD\_RES 82 82  
 FT CONFLICT 49 50  
 SQ SEQUENCE 84 AA; 9436 MW; D75CABE7 CRC32;

Query Match 17.2%; Score 5; DB 1; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 8.42e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 VDKNT 26  
 QY 24 VDKNT 28

Search completed: Sat Aug 28 14:51:54 1999  
 Job time : 44 secs.



94	5	17.2	216	2	06547	HYPOTHETICAL 23.4 KD P	2.03e+02	167	5	17.2	339	4	076071	WD40 PROTEIN CIAO 1.	2.03e+02
95	5	17.2	217	11	08664	BCL7C MRNA.	2.03e+02	168	5	17.2	341	2	007863	RPKA.	2.03e+02
96	5	17.2	219	2	067826	HYPOTHETICAL 25.3 KD P	2.03e+02	169	5	17.2	342	4	043700	CD33L2.	2.03e+02
97	5	17.2	223	5	P91245	CODED FOR BY C. ELEGAN	2.03e+02	170	5	17.2	342	10	023124	F19G10.6 PROTEIN.	2.03e+02
98	5	17.2	225	10	Q40129	HYPOTHETICAL 25.2 KD P	2.03e+02	171	5	17.2	344	5	002102	COSMID C18E3.	2.03e+02
99	5	17.2	226	4	Q16087	SMOOTH MUSCLE MYOSIN H	2.03e+02	172	5	17.2	347	10	082661	SUCCINYL-COA-LIGASE AL	2.03e+02
100	5	17.2	226	8	Q79551	ATP SYNTHASE A CHAIN (	2.03e+02	173	5	17.2	349	2	056920	TRSH.	2.03e+02
101	5	17.2	229	2	067801	HYPOTHETICAL 25.6 KD P	2.03e+02	174	5	17.2	351	3	042781	MITOGEN-ACTIVATED PROT	2.03e+02
102	5	17.2	234	14	Q92444	ACMPV ORF81.	2.03e+02	175	5	17.2	352	14	083053	NSI GENE (FRAGMENT).	2.03e+02
103	5	17.2	234	10	Q39559	SITE-SPECIFIC DNA ENDO	2.03e+02	176	5	17.2	354	2	034538	YCCA.	2.03e+02
104	5	17.2	234	4	060443	ICERE-1.	2.03e+02	177	5	17.2	355	3	000859	MITOGEN-ACTIVATED PROT	2.03e+02
105	5	17.2	234	5	Q17006	T2B3E12.1 PROTEIN.	2.03e+02	178	5	17.2	355	1	029783	IMMUNOGENIC PROTEIN (B	2.03e+02
106	5	17.2	234	2	053327	MEMBRANE FUSION PROTEI	2.03e+02	179	5	17.2	356	3	092246	PATHOGENICITY MAP KINA	2.03e+02
107	5	17.2	235	11	Q63068	CYTOSOLIC RESINIFERATO	2.03e+02	180	5	17.2	356	3	060123	CARBOXYLESTERASE (EC 3	2.03e+02
108	5	17.2	237	2	Q4923	HYPOTHETICAL 26.1 KD P	2.03e+02	181	5	17.2	357	10	040344	FLAVONONE-3-HYDROXYLAS	2.03e+02
109	5	17.2	237	5	Q17242	MICROFILAMENT SHEATH P	2.03e+02	182	5	17.2	357	10	042927	NARINGENIN 3-DIOXYGENA	2.03e+02
110	5	17.2	239	2	087623	PUTATIVE SUBSTRATE BIN	2.03e+02	183	5	17.2	358	6	Q29134	VON WILLEBRAND FACTOR	2.03e+02
111	5	17.2	240	2	P73861	HYPOTHETICAL 26.2 KD P	2.03e+02	184	5	17.2	358	8	033545	RIBULOSE-1,5-BISPHOSPH	2.03e+02
112	5	17.2	241	2	068961	PUTATIVE MEMBRANE FUSI	2.03e+02	185	5	17.2	360	10	023593	HYPOTHETICAL 38.9 KD P	2.03e+02
113	5	17.2	243	1	Q30149	NITRATE ABC TRANSPORTE	2.03e+02	186	5	17.2	361	3	074191	EKBP39P.	2.03e+02
114	5	17.2	245	1	Q58723	HYPOTHETICAL PROTEIN M	2.03e+02	187	5	17.2	361	1	057936	361AA LONG HYPOTHETICA	2.03e+02
115	5	17.2	246	2	P72704	PROBABLE PEPTIDYL-PROL	2.03e+02	188	5	17.2	362	1	027988	CYTOCHROME OXIDASE.SUB	2.03e+02
116	5	17.2	249	8	Q53300	NADH DEHYDROGENASE SUB	2.03e+02	189	5	17.2	364	6	Q28129	INTERSTITIAL RETINOL B	2.03e+02
117	5	17.2	253	2	P74526	HYPOTHETICAL 28.7 KD P	2.03e+02	190	5	17.2	366	10	027410	EWAM (ACTIN-MODULATOR	2.03e+02
118	5	17.2	255	4	Q4876	MUC1 MRNA.	2.03e+02	191	5	17.2	369	10	040754	FLAVANONE 3-HYDROXYLAS	2.03e+02
119	5	17.2	255	13	Q73747	FRG1.	2.03e+02	192	5	17.2	369	5	020804	F55C10.1 PROTEIN.	2.03e+02
120	5	17.2	256	2	083899	HYPOTHETICAL 28.2 KD P	2.03e+02	193	5	17.2	371	14	086843	FIBER.	2.03e+02
121	5	17.2	257	1	Q58841	HYPOTHETICAL PROTEIN M	2.03e+02	194	5	17.2	374	14	066210	HYPOTHETICAL 40.1 KD P	2.03e+02
122	5	17.2	265	10	080699	F8K4.13 PROTEIN.	2.03e+02	195	5	17.2	374	11	061183	POLY(A) POLYMERASE (EC	2.03e+02
123	5	17.2	269	1	Q92908	HYPOTHETICAL 30.2 KD P	2.03e+02	196	5	17.2	374	14	098583	PECV-1 UNIDENTIFIED OR	2.03e+02
124	5	17.2	273	10	Q39028	ATMYB2.	2.03e+02	197	5	17.2	376	5	094301	SIMILAR TO UBIQUITIN T	2.03e+02
125	5	17.2	276	10	Q43295	APS KINASE PRECURSOR (	2.03e+02	198	5	17.2	376	5	076919	EG-9D2.2 PROTEIN.	2.03e+02
126	5	17.2	279	10	Q14543	GAMMA-GLIADIN PRECURSO	2.03e+02	199	5	17.2	386	14	087437	ENVELOPE GLYCOPROTEIN	2.03e+02
127	5	17.2	280	1	Q92730	CONSERVED HYPOTHETICAL	2.03e+02	200	5	17.2	387	4	043179	TUMOR-SUPPRESSING SUBC	2.03e+02
128	5	17.2	280	2	P96043	HYPOTHETICAL 31.7 KD P	2.03e+02	201	5	17.2	387	11	060863	VON WILLEBRAND FACTOR	2.03e+02
129	5	17.2	280	10	Q41602	GAMMA GLIADIN (FRAGMEN	2.03e+02	202	5	17.2	392	2	P72924	HYPOTHETICAL 43.0 KD P	2.03e+02
130	5	17.2	282	2	P77068	INTIMIN (FRAGMENT).	2.03e+02	203	5	17.2	392	3	066205	CHROMOSOME XII COSMID	2.03e+02
131	5	17.2	282	2	P77628	ITIMIN (FRAGMENT).	2.03e+02	204	5	17.2	393	11	088371	SEMAPHORIN L (FRAGMENT	2.03e+02
132	5	17.2	286	2	P77739	FROM BASES 1798041 TO	2.03e+02	205	5	17.2	395	2	055147	HYPOTHETICAL 43.1 KD P	2.03e+02
133	5	17.2	287	14	Q85076	CAPSID PROTEIN (FRAGME	2.03e+02	206	5	17.2	400	14	071958	POLYPROTEIN (FRAGMENT)	2.03e+02
134	5	17.2	287	10	Q85087	PYROVATE DEHYDROGENASE	2.03e+02	207	5	17.2	400	8	Q37619	NADH DEHYDROGENASE (UB	2.03e+02
135	5	17.2	288	2	054193	2-HYDROXY-6-KETONONA-2	2.03e+02	208	5	17.2	405	2	088024	PUTATIVE TWO-COMPONENT	2.03e+02
136	5	17.2	288	5	Q44093	ADP/ATP TRANSLOCASE (F	2.03e+02	209	5	17.2	406	6	077672	VON WILLEBRAND FACTOR	2.03e+02
137	5	17.2	288	5	Q44094	ADP/ATP TRANSLOCASE (F	2.03e+02	210	5	17.2	407	2	P96559	GLYCOSYLTRANSFERASE GT	2.03e+02
138	5	17.2	289	2	P73086	ABC TRANSPORTER.	2.03e+02	211	5	17.2	408	2	005898	MANA.	2.03e+02
139	5	17.2	290	10	Q65851	XANTHI ARCA 3 (FRAGMEN	2.03e+02	212	5	17.2	409	5	P91586	COS41.8.	2.03e+02
140	5	17.2	290	5	Q94482	ORF DGL003.	2.03e+02	213	5	17.2	410	2	P72454	L-ALANINE:N-AMIDINO-3-	2.03e+02
141	5	17.2	298	2	Q25832	CONSERVED HYPOTHETICAL	2.03e+02	214	5	17.2	412	1	027205	POLYFERREDOXIN (MVHB).	2.03e+02
142	5	17.2	299	2	Q87564	NATA.	2.03e+02	215	5	17.2	412	1	050784	FERREDOXIN.	2.03e+02
143	5	17.2	304	10	Q49745	R2R3-MYB TRANSCRIPTION	2.03e+02	216	5	17.2	416	2	084763	UDP-N-ACETYLMURAMOYLAL	2.03e+02
144	5	17.2	304	5	Q18176	W05G6.1.	2.03e+02	217	5	17.2	424	14	069972	GP120 (FRAGMENT).	2.03e+02
145	5	17.2	304	5	Q20981	F58E6.5. PROTEIN.	2.03e+02	218	5	17.2	425	6	Q28759	VON WILLEBRAND FACTOR	2.03e+02
146	5	17.2	304	10	Q39155	MYB-RELATED PROTEIN.	2.03e+02	219	5	17.2	425	14	Q64796	EARLY 49 KDA PROTEIN.	2.03e+02
147	5	17.2	308	3	Q12010	CHROMOSOME XV READING	2.03e+02	220	5	17.2	429	5	Q17472	COSMID B0303 (FRAGMENT	2.03e+02
148	5	17.2	310	2	Q52556	ESTERASE.	2.03e+02	221	5	17.2	429	2	052702	APALI METHYLTRANSFERAS	2.03e+02
149	5	17.2	311	10	Q23381	SIMILARITY TO UTP-GLUC	2.03e+02	222	5	17.2	429	2	P70750	APALI METHYLASE.	2.03e+02
150	5	17.2	312	5	Q17557	SIMILARITY TO S. CEREVISI	2.03e+02	223	5	17.2	430	3	001823	ORF-3.	2.03e+02
151	5	17.2	313	5	Q20166	F38E11.3 PROTEIN.	2.03e+02	224	5	17.2	432	10	082460	RAC GTPASE ACTIVATING	2.03e+02
152	5	17.2	315	5	P91282	COSMID F27C1.	2.03e+02	225	5	17.2	433	8	Q32849	RIBULOSE BISPHOSPHATE	2.03e+02
153	5	17.2	317	9	Q80052	ORF 13.	2.03e+02	226	5	17.2	433	5	045500	F39B2.8 PROTEIN.	2.03e+02
154	5	17.2	317	14	Q66606	VIRAL OXYPOLYPROTEIN.	2.03e+02	227	5	17.2	438	8	Q33409	RIBULOSE BISPHOSPHATE	2.03e+02
155	5	17.2	320	2	P94346	HEME O OXYGENASE.	2.03e+02	228	5	17.2	438	1	058314	438AA LONG HYPOTHETICA	2.03e+02
156	5	17.2	323	2	Q52354	LACTATE DEHYDROGENASE.	2.03e+02	229	5	17.2	440	4	015388	OB BINDING PROTEIN-1.	2.03e+02
157	5	17.2	324	4	Q16107	C6.1A.	2.03e+02	230	5	17.2	441	2	085376	ZAPC1.	2.03e+02
158	5	17.2	324	2	P73405	PYROVATE DEHYDROGENASE	2.03e+02	231	5	17.2	442	4	043699	CD33L1.	2.03e+02
159	5	17.2	327	3	Q13926	HYPOTHETICAL 37.8 KD P	2.03e+02	232	5	17.2	443	5	061388	DELTA6-FATTY-ACID-DESA	2.03e+02
160	5	17.2	328	4	Q16615	EPISIALIN VARIANT B (F	2.03e+02	233	5	17.2	445	11	009173	HOMOGENITISATE 1,2-DIOX	2.03e+02
161	5	17.2	328	10	Q40403	G PROTEIN BETA-SUBUNIT	2.03e+02	234	5	17.2	447	2	P76097	FROM BASES 1485094 TO	2.03e+02
162	5	17.2	329	5	Q44878	T24E12.6 PROTEIN.	2.03e+02	235	5	17.2	447	4	Q13133	NUCLEAR ORPHAN RECEPT	2.03e+02
163	5	17.2	334	14	Q83150	ORF H55.	2.03e+02	236	5	17.2	448	8	Q32263	RIBULOSE-1,5-BISPHOSPH	2.03e+02
164	5	17.2	336	2	Q31619	YJBU PROTEIN.	2.03e+02	237	5	17.2	449	5	062027	C04H5.5 PROTEIN.	2.03e+02
165	5	17.2	338	5	Q24795	CALCIUM-BINDING PROTEI	2.03e+02	238	5	17.2	451	2	P71003	HYPOTHETICAL 49.4 KD P	2.03e+02
166	5	17.2	339	3	Q14024	HYPOTHETICAL 39.9 KD P	2.03e+02	239	5	17.2	451	10	082690	MAT1.	2.03e+02

240	5	17.2	452	8	Q33100	RIBULOSE BIPHOSPHATE	2.03e+02	313	5	17.2	591	4	Q13041	P67.	2.03e+02
241	5	17.2	452	10	Q23360	GLUCOSYLTRANSFERASE.	2.03e+02	314	5	17.2	592	2	O53538	HYPOTHETICAL 59.8 KD P	2.03e+02
242	5	17.2	453	14	O84638	GENOME, PARTIAL SEQUEN	2.03e+02	315	5	17.2	594	10	O48590	GTL1 PROTEIN.	2.03e+02
243	5	17.2	454	10	O22388	T27A16.23 PROTEIN.	2.03e+02	316	5	17.2	598	14	O56653	MAJOR COAT PROTEIN VP2	2.03e+02
244	5	17.2	458	4	O60326	KTAA0582 PROTEIN (FRAG	2.03e+02	317	5	17.2	606	2	O13566	YF1X PROTEIN.	2.03e+02
245	5	17.2	460	2	O06249	GADB.	2.03e+02	318	5	17.2	610	2	O52961	YF1X.	2.03e+02
246	5	17.2	460	2	O34448	YOPQ PROTEIN.	2.03e+02	319	5	17.2	610	2	O51229	DNA MISMATCH REPAIR PR	2.03e+02
247	5	17.2	460	9	O64100	HYPOTHETICAL 53.5 KD P	2.03e+02	320	5	17.2	611	5	O20952	F57F5.4 PROTEIN.	2.03e+02
248	5	17.2	463	2	P95224	HYPOTHETICAL 49.1 KD P	2.03e+02	321	5	17.2	612	5	O44106	GLUCOSE DEHYDROGENASE	2.03e+02
249	5	17.2	464	8	O32558	RIBULOSE BIPHOSPHATE	2.03e+02	322	5	17.2	622	5	O24043	HYPOTHETICAL PROTEIN (	2.03e+02
250	5	17.2	466	2	O48801	ICMX.	2.03e+02	323	5	17.2	622	5	Q93602	F26D10.4.	2.03e+02
251	5	17.2	466	8	O32144	RIBULOSE BIPHOSPHATE	2.03e+02	324	5	17.2	629	5	P90887	F55A11.6 PROTEIN.	2.03e+02
252	5	17.2	467	3	Q06456	CHROMOSOME XVI COSMID	2.03e+02	325	5	17.2	636	3	O60126	DNA TOPOISOMERASE III.	2.03e+02
253	5	17.2	467	3	Q04930	HYPOTHETICAL 51.7 KD P	2.03e+02	326	5	17.2	640	2	O54152	TRANSFERASE.	2.03e+02
254	5	17.2	468	5	Q93779	F53H4.5 PROTEIN.	2.03e+02	327	5	17.2	643	3	O04177	D9509.17P.	2.03e+02
255	5	17.2	468	14	P88838	NEURAMINIDASE.	2.03e+02	328	5	17.2	645	3	O42946	KINASE-BINDING PROTEIN	2.03e+02
256	5	17.2	469	14	P88837	NEURAMINIDASE.	2.03e+02	329	5	17.2	645	3	O42946	HORMONE RECEPTOR 3C.	2.03e+02
257	5	17.2	473	5	P23221	W08D2.4 PROTEIN.	2.03e+02	330	5	17.2	651	5	O02643	COMPLEMENT H FACTOR (F	2.03e+02
258	5	17.2	474	10	O48591	GIL2 GENE.	2.03e+02	331	5	17.2	657	4	O14006	MAJOR OUTER ENVELOPE G	2.03e+02
259	5	17.2	475	4	O00756	LACTOFERRIN (FRAGMENT)	2.03e+02	332	5	17.2	658	14	O66536	KIAA0405.	2.03e+02
260	5	17.2	476	14	O92494	ACMNPV ORF142.	2.03e+02	333	5	17.2	660	4	O43155	PEPTIDYLARGININE DEIMI	2.03e+02
261	5	17.2	476	2	O52892	NOEA.	2.03e+02	334	5	17.2	666	11	O35117	PEPTIDYLARGININE L.	2.03e+02
262	5	17.2	477	5	O19580	F18H3.4 PROTEIN.	2.03e+02	335	5	17.2	666	4	O75326	PEPTIDYLARGININE DEIMI	2.03e+02
263	5	17.2	478	10	O23382	SIMILARITY TO UTP-GLUC	2.03e+02	336	5	17.2	666	11	O88807	CCP MODULES 3-12 WITH	2.03e+02
264	5	17.2	479	2	P71309	HYPOTHETICAL 50.4 KD P	2.03e+02	337	5	17.2	669	6	O28085	SIMILAR TO A HUMAN PUT	2.03e+02
265	5	17.2	479	3	Q03942	HYPOTHETICAL 55.9 KD P	2.03e+02	338	5	17.2	678	10	O65709	HYPOTHETICAL 74.0 KD P	2.03e+02
266	5	17.2	481	5	O94410	ZK678.5 PROTEIN.	2.03e+02	339	5	17.2	683	4	O43387	FM315.	2.03e+02
267	5	17.2	482	2	O55698	HYPOTHETICAL 50.8 KD P	2.03e+02	340	5	17.2	683	5	O76951	PROPHENOXIDASE.	2.03e+02
268	5	17.2	495	10	O04375	B' REGULATORY SUBUNIT	2.03e+02	341	5	17.2	683	4	O14978	ZINC FINGER PROTEIN 96	2.03e+02
269	5	17.2	495	10	O23330	HYPOTHETICAL 53.6 KD P	2.03e+02	342	5	17.2	689	5	O94363	W10D5.3A PROTEIN.	2.03e+02
270	5	17.2	497	2	O55182	O-SUCCINYL BENZOIC ACID	2.03e+02	343	5	17.2	691	13	O57658	BONE MORPHOGENETIC PRO	2.03e+02
271	5	17.2	501	2	O55762	HYPOTHETICAL 52.4 KD P	2.03e+02	344	5	17.2	705	4	O43377	BAC CLONE RGL18D07 FRO	2.03e+02
272	5	17.2	504	13	Q93304	CYTOCHROME P450 MONOOX	2.03e+02	345	5	17.2	706	4	O16780	HLF2-ENCODED PROTEIN P	2.03e+02
273	5	17.2	504	13	Q93297	CYTOCHROME P450.	2.03e+02	346	5	17.2	707	3	O13373	M ANTIGEN.	2.03e+02
274	5	17.2	505	11	O60569	PROCOLLAGEN, TYPE XVII	2.03e+02	347	5	17.2	708	2	O46444	QUINOAEMOPROTEIN ETHA	2.03e+02
275	5	17.2	508	4	O43558	EGFR SIGNALLING ADAPTO	2.03e+02	348	5	17.2	709	5	O16783	COSMID T21D12.	2.03e+02
276	5	17.2	509	5	O09483	HYPOTHETICAL 50.5 KD P	2.03e+02	349	5	17.2	710	4	O16785	NEUTROPHIL LACTOFERRIN	2.03e+02
277	5	17.2	509	5	O01380	UNC-18 (FRAGMENT).	2.03e+02	350	5	17.2	711	4	O16786	LACTOFERRIN.	2.03e+02
278	5	17.2	509	10	O96383	ASPARTIC PROTEINASE.	2.03e+02	351	5	17.2	711	4	O16789	LACTOFERRIN PRECURSOR.	2.03e+02
279	5	17.2	512	2	O45372	FRUCTOSYLTRANSFERASE (	2.03e+02	352	5	17.2	712	13	O42443	TAP2.	2.03e+02
280	5	17.2	515	14	O89233	NUCLEOCAPSID PROTEIN.	2.03e+02	353	5	17.2	715	10	O23625	AB13 PROTEIN.	2.03e+02
281	5	17.2	515	14	O81075	NUCLEOCAPSID PROTEIN.	2.03e+02	354	5	17.2	720	3	O13878	HYPOTHETICAL 79.4 KD P	2.03e+02
282	5	17.2	515	9	O64283	HYPOTHETICAL 58.5 KD P	2.03e+02	355	5	17.2	725	13	O57521	HEAT SHOCK PROTEIN HSP	2.03e+02
283	5	17.2	520	5	O76247	TRANSCRIPTION FACTOR Y	2.03e+02	356	5	17.2	729	14	O83430	MAJOR COAT PROTEIN VP1	2.03e+02
284	5	17.2	521	6	O46651	ALCAM, PARTIAL (FRAGME	2.03e+02	357	5	17.2	735	14	O56652	MAJOR COAT PROTEIN VP1.	2.03e+02
285	5	17.2	521	6	O46634	ALCAM, PARTIAL (FRAGME	2.03e+02	358	5	17.2	736	14	O56137	PENICILLIN-BINDING PRO	2.03e+02
286	5	17.2	525	5	O44079	CHITINASE.	2.03e+02	359	5	17.2	742	2	O07106	B24 OCYTE NUCLEAR PRO	2.03e+02
287	5	17.2	531	2	O50339	PUTATIVE LIPOPROTEIN.	2.03e+02	360	5	17.2	744	13	Q91149	YURC PROTEIN.	2.03e+02
288	5	17.2	533	14	O92917	MAJOR COAT PROTEIN VP3	2.03e+02	361	5	17.2	745	2	O32144	R04B5.9 PROTEIN.	2.03e+02
289	5	17.2	535	5	O17595	SIMILARITY TO MYOSIN H	2.03e+02	362	5	17.2	745	5	O21706	CHEA.	2.03e+02
290	5	17.2	538	11	O60525	GUANINE NUCLEOTIDE-BIN	2.03e+02	363	5	17.2	747	2	O52261	CHEA.	2.03e+02
291	5	17.2	544	5	O22496	T14G10.7 PROTEIN.	2.03e+02	364	5	17.2	753	2	O87124	CHEA.	2.03e+02
292	5	17.2	546	5	O27547	HORMONE RECEPTOR 3.	2.03e+02	365	5	17.2	754	4	O60672	CELL CYCLE REGULATORY	2.03e+02
293	5	17.2	547	5	O18609	SIMILAR TO YEAST VACUO	2.03e+02	366	5	17.2	754	4	O60934	NEBRIN.	2.03e+02
294	5	17.2	547	11	P97881	PUTATIVE CELL SURFACE	2.03e+02	367	5	17.2	756	6	O28483	TMDC II MRNA.	2.03e+02
295	5	17.2	557	5	O54129	M75 PROTEIN.	2.03e+02	368	5	17.2	759	10	O23135	HYPOTHETICAL PROTEIN.	2.03e+02
296	5	17.2	558	1	O98131	558AA LONG HYPOTHETICA	2.03e+02	369	5	17.2	767	5	O76977	HYPERPOLARIZATION-ACTI	2.03e+02
297	5	17.2	565	2	O24929	METHYL-ACCEPTING CEMO	2.03e+02	370	5	17.2	768	4	O60279	KIAA0527 PROTEIN (FRAG	2.03e+02
298	5	17.2	565	2	P75472	F10-ORF565 PROTEIN.	2.03e+02	371	5	17.2	772	5	O94361	W10D5.3C PROTEIN.	2.03e+02
299	5	17.2	568	4	O13647	TESTIS SPECIFIC BASIC	2.03e+02	372	5	17.2	773	5	O76871	EG:100G7.2 PROTEIN.	2.03e+02
300	5	17.2	568	5	O27212	ACTINULIN P60.	2.03e+02	373	5	17.2	777	10	O81868	HYPOTHETICAL 87.4 KD P	2.03e+02
301	5	17.2	571	2	O84959	DEOXYRIBODIPYRIMIDINE	2.03e+02	374	5	17.2	780	4	O75157	KIAA0669 PROTEIN.	2.03e+02
302	5	17.2	574	5	O22805	CODED FOR BY C. ELEGAN	2.03e+02	375	5	17.2	781	2	O52753	XYNB PRECURSOR.	2.03e+02
303	5	17.2	574	5	O23461	SIMILAR TO BREAKPOINT	2.03e+02	376	5	17.2	784	11	O35542	CADHERIN-5.	2.03e+02
304	5	17.2	575	5	O62223	F35E2.5 PROTEIN.	2.03e+02	377	5	17.2	791	5	O24042	HYPOTHETICAL 89.2 KD P	2.03e+02
305	5	17.2	575	10	O23083	SIMILARITY TO DNA BIND	2.03e+02	378	5	17.2	794	2	P75493	PUTATIVE LIPOPROTEIN.	2.03e+02
306	5	17.2	577	4	O60907	TRANSUDIN (BETA) LIKE	2.03e+02	379	5	17.2	806	5	O17851	COSMID C09B9.	2.03e+02
307	5	17.2	578	3	O74722	PYRUVATE DECARBOXYLASE	2.03e+02	380	5	17.2	811	5	P91482	SIMILARITY TO C3HC4-TY	2.03e+02
308	5	17.2	578	5	O24238	ALKALINE PHOSPHATASE 4	2.03e+02	381	5	17.2	814	2	O52984	HYPOTHETICAL 90.3 KD P	2.03e+02
309	5	17.2	582	4	O60892	MEMD PROTEIN (FRAGMENT	2.03e+02	382	5	17.2	814	10	O65715	HYPOTHETICAL 91.9 KD P	2.03e+02
310	5	17.2	583	5	O27472	SIMILAR TO CYTOCHROME	2.03e+02	383	5	17.2	815	10	O41553	TRITICUM AESTIVUM HIG	2.03e+02
311	5	17.2	583	11	O70136	ACTIVATED LEUKOCYTE CE	2.03e+02	384	5	17.2	815	13	P87376	B24 PROTEIN.	2.03e+02
312	5	17.2	583	11	O35112	HB2, COMPLETE CDS.	2.03e+02	385	5	17.2	821	2	O52902	HELO.	2.03e+02

386	5	17.2	828	2	O51843	2.03e+02	PROBABLY METHYLASE/HEL	2.03e+02
387	5	17.2	828	5	O20550	2.03e+02	CUT-LIKE 2 (CUX-2).	2.03e+02
388	5	17.2	831	2	O34796	2.03e+02	COUNTERPART OF HSV-1 G	2.03e+02
389	5	17.2	832	4	O73427	2.03e+02	ZC8. 6 PROTEIN.	2.03e+02
390	5	17.2	836	5	O77347	2.03e+02	LEUCIN RICH NEURONAL P	2.03e+02
391	5	17.2	845	11	O35833	2.03e+02	MAL3P4.11 PROTEIN.	2.03e+02
392	5	17.2	846	4	O15516	2.03e+02	MATRN 3.	2.03e+02
393	5	17.2	850	2	P77075	2.03e+02	CLOCK.	2.03e+02
394	5	17.2	850	4	O00234	2.03e+02	INTIMIN (FRAGMENT).	2.03e+02
395	5	17.2	856	14	O56854	2.03e+02	KIAA0235 (FRAGMENT).	2.03e+02
396	5	17.2	861	6	O29470	2.03e+02	GP340.	2.03e+02
397	5	17.2	864	14	O69408	2.03e+02	CGMP-PHOSPHODIESTERASE	2.03e+02
398	5	17.2	872	5	P90523	2.03e+02	GB HOMOLOG.	2.03e+02
399	5	17.2	873	5	O20072	2.03e+02	PUTATIVE TRANSCRIPTION	2.03e+02
400	5	17.2	876	10	O64483	2.03e+02	SIMILAR TO VAPPAE 116	2.03e+02
401	5	17.2	876	10	O64537	2.03e+02	HYPOTHETICAL 98.7 KD P	2.03e+02
402	5	17.2	907	13	O68537	2.03e+02	MAJOR OUTER ENVELOPE G	2.03e+02
403	5	17.2	907	13	O98850	2.03e+02	NEURONAL MYOSIN LIGHT	2.03e+02
404	5	17.2	911	2	O48044	2.03e+02	TRANSFERRIN BINDING PR	2.03e+02
405	5	17.2	912	2	O48037	2.03e+02	TRANSFERRIN BINDING PR	2.03e+02
406	5	17.2	913	2	O48035	2.03e+02	TRANSFERRIN BINDING PR	2.03e+02
407	5	17.2	913	11	O88826	2.03e+02	GOB-5 PROTEIN.	2.03e+02
408	5	17.2	914	2	O48039	2.03e+02	TRANSFERRIN BINDING PR	2.03e+02
409	5	17.2	922	13	O93405	2.03e+02	FIBRONECTIN.	2.03e+02
410	5	17.2	924	5	O20189	2.03e+02	F39D8.1C PROTEIN.	2.03e+02
411	5	17.2	926	3	O13305	2.03e+02	PROTEASE 1.	2.03e+02
412	5	17.2	930	2	O54366	2.03e+02	TRANSFERRIN BINDING PR	2.03e+02
413	5	17.2	930	11	O61143	2.03e+02	CALCIUM ENTRY CHANNEL.	2.03e+02
414	5	17.2	935	4	O13078	2.03e+02	ARI (TCP20) (FRAGMENT)	2.03e+02
415	5	17.2	938	2	O52620	2.03e+02	INTIMIN.	2.03e+02
416	5	17.2	939	2	O57379	2.03e+02	INTIMIN.	2.03e+02
417	5	17.2	939	2	P77067	2.03e+02	INTIMIN.	2.03e+02
418	5	17.2	952	5	O48446	2.03e+02	TO4B8.5 PROTEIN.	2.03e+02
419	5	17.2	953	3	O01869	2.03e+02	HYPOTHETICAL 109.5 KD	2.03e+02
420	5	17.2	960	5	OQ5507	2.03e+02	FL-160-3 FLAGELLUM-ASS	2.03e+02
421	5	17.2	966	5	O44066	2.03e+02	DNA POLYMERASE ALPHA (	2.03e+02
422	5	17.2	972	5	O40938	2.03e+02	THROMBOSPONDIN-LIKE PR	2.03e+02
423	5	17.2	979	2	O53609	2.03e+02	PUTATIVE MEMBRANE PROT	2.03e+02
424	5	17.2	995	5	O46091	2.03e+02	1-EVIDENCE-PREDICTED B	2.03e+02
425	5	17.2	1003	5	OQ9060	2.03e+02	FL-160-2 FLAGELLUM-ASS	2.03e+02
426	5	17.2	1003	5	O26889	2.03e+02	COMPLEMENT REGULATORY	2.03e+02
427	5	17.2	1014	6	O77813	2.03e+02	NEBULIN (FRAGMENT).	2.03e+02
428	5	17.2	1027	14	O77373	2.03e+02	POL POLYPROTEIN (FRAGM	2.03e+02
429	5	17.2	1030	12	O09266	2.03e+02	BETA-GALACTOSIDASE.	2.03e+02
430	5	17.2	1034	3	O60146	2.03e+02	HYPOTHETICAL 116.4 KD	2.03e+02
431	5	17.2	1035	12	O09267	2.03e+02	BETA-GALACTOSIDASE.	2.03e+02
432	5	17.2	1043	5	O17644	2.03e+02	C34F6.1 PROTEIN.	2.03e+02
433	5	17.2	1056	2	P97096	2.03e+02	CLOVING VECTOR PLAC21,	2.03e+02
434	5	17.2	1063	10	O80596	2.03e+02	T2711.7 PROTEIN.	2.03e+02
435	5	17.2	1086	10	O49369	2.03e+02	HYPOTHETICAL 122.7 KD	2.03e+02
436	5	17.2	1130	14	O88282	2.03e+02	ENVELOPE PROTEIN.	2.03e+02
437	5	17.2	1141	10	O65598	2.03e+02	HYPOTHETICAL 129.7 KD	2.03e+02
438	5	17.2	1144	2	OQ0371	2.03e+02	PI-LIKE ADHESIN.	2.03e+02
439	5	17.2	1151	13	O57580	2.03e+02	HIGH MOLECULAR MASS NU	2.03e+02
440	5	17.2	1155	3	O14239	2.03e+02	HYPOTHETICAL 128.3 KD	2.03e+02
441	5	17.2	1175	5	O17732	2.03e+02	D2023.2 PROTEIN.	2.03e+02
442	5	17.2	1186	7	P79457	2.03e+02	UBIQUITOUSLY TRANSCRIB	2.03e+02
443	5	17.2	1198	3	O13353	2.03e+02	CLASS IV CHITIN SYNTHA	2.03e+02
444	5	17.2	1219	13	O73749	2.03e+02	NUCLEOPORIN NUP153 HOM	2.03e+02
445	5	17.2	1225	5	O15784	2.03e+02	HISTIDINE KINASE C.	2.03e+02
446	5	17.2	1232	13	O90284	2.03e+02	L1-LIKE CELL ADHESION	2.03e+02
447	5	17.2	1246	14	O36977	2.03e+02	POL POLYPROTEIN (FRAGM	2.03e+02
448	5	17.2	1252	11	O70470	2.03e+02	CORTACTIN-BINDING PROT	2.03e+02
449	5	17.2	1257	2	P96746	2.03e+02	2-OXOGUTARATE DEHYDRO	2.03e+02
450	5	17.2	1266	5	OQ0331	2.03e+02	PGP-4 PROTEIN.	2.03e+02
451	5	17.2	1268	5	OQ0335	2.03e+02	PGP-3 PROTEIN.	2.03e+02
452	5	17.2	1277	13	O98902	2.03e+02	NEURAL CELL ADHESION M	2.03e+02
453	5	17.2	1285	11	O70305	2.03e+02	SPINOCEREBELLAR ATAXIA	2.03e+02
454	5	17.2	1287	13	O93437	2.03e+02	ABC TRANSPORTER PROTEI	2.03e+02
455	5	17.2	1290	2	O48756	2.03e+02	DEXTRANSUCRASE.	2.03e+02
456	5	17.2	1293	10	O04251	2.03e+02	PREDICTED PROTEIN OF U	2.03e+02
457	5	17.2	1312	4	O93700	2.03e+02	ATAXIN-2.	2.03e+02
458	5	17.2	1327	5	O61072	2.03e+02	GUANIDINE NUCLEOTIDE E	2.03e+02
459	5	17.2	1327	5	O61072	2.03e+02	SIMILAR TO THE PROTEIN	2.03e+02
460	5	17.2	1398	5	P91099	2.03e+02		2.03e+02
461	5	17.2	1426	11	P70298	2.03e+02	PROBABLY METHYLASE/HEL	2.03e+02
462	5	17.2	1442	14	O42066	2.03e+02	COUNTERPART OF HSV-1 G	2.03e+02
463	5	17.2	1474	5	O62504	2.03e+02	ZC8. 6 PROTEIN.	2.03e+02
464	5	17.2	1476	13	O90285	2.03e+02	PUTATIVE EXTRACELLULAR	2.03e+02
465	5	17.2	1624	2	O53203	2.03e+02	HYPOTHETICAL 176.9 KD	2.03e+02
466	5	17.2	1624	5	O45000	2.03e+02	W03D8.6 PROTEIN.	2.03e+02
467	5	17.2	1662	10	O81630	2.03e+02	F8M12.22 PROTEIN.	2.03e+02
468	5	17.2	1677	11	O70373	2.03e+02	XIN	2.03e+02
469	5	17.2	1716	4	O14528	2.03e+02	KIAA0292 (FRAGMENT).	2.03e+02
470	5	17.2	1736	14	O92808	2.03e+02	PR180.	2.03e+02
471	5	17.2	1740	1	P54844	2.03e+02	RIBONUCLEOTIDE REDUCTA	2.03e+02
472	5	17.2	1742	2	O55583	2.03e+02	HYPOTHETICAL 179.7 KD	2.03e+02
473	5	17.2	1774	11	O62001	2.03e+02	PROCOLLAGEN, TYPE XVII	2.03e+02
474	5	17.2	1800	3	P87162	2.03e+02	6-METHYLSALICYLIC ACID	2.03e+02
475	5	17.2	1803	3	OQ0290	2.03e+02	POLYKETIDE SYNTHASE.	2.03e+02
476	5	17.2	1807	3	O13661	2.03e+02	HYPOTHETICAL 229.9KD P	2.03e+02
477	5	17.2	1813	6	O62653	2.03e+02	KUSPASE-ISOMALTASE.	2.03e+02
478	5	17.2	1903	3	P87204	2.03e+02	FKBP PROTEIN INVOLVED	2.03e+02
479	5	17.2	1905	3	O92225	2.03e+02	1,3'-BETA-D-GLUCAN SYNT	2.03e+02
480	5	17.2	1919	3	O42998	2.03e+02	HYPOTHETICAL 217.7 KD	2.03e+02
481	5	17.2	1967	14	O10378	2.03e+02	LARGE PROTEIN	2.03e+02
482	5	17.2	1999	11	O83731	2.03e+02	NEURONAL MYOSIN HEAVY	2.03e+02
483	5	17.2	2034	5	O17388	2.03e+02	FER-1.	2.03e+02
484	5	17.2	2129	2	O85020	2.03e+02	FXBC (FRAGMENT).	2.03e+02
485	5	17.2	2228	5	O60991	2.03e+02	ERYTHROCYTE MEMBRANE P	2.03e+02
486	5	17.2	2229	5	O19853	2.03e+02	SIMILAR TO EGF-LIKE RE	2.03e+02
487	5	17.2	2250	11	O93406	2.03e+02	FIBRONECTIN.	2.03e+02
488	5	17.2	2505	11	O83577	2.03e+02	FATTY ACID SYNTHASE.	2.03e+02
489	5	17.2	2567	13	O90574	2.03e+02	FILAMIN.	2.03e+02
490	5	17.2	2602	4	O75369	2.03e+02	BETA-FILAMIN.	2.03e+02
491	5	17.2	2723	2	O30479	2.03e+02	PKS MODULE 1.	2.03e+02
492	5	17.2	3007	4	O14215	2.03e+02	NEBULIN (FRAGMENT).	2.03e+02
493	5	17.2	3036	5	O17575	2.03e+02	CO898.7.	2.03e+02
494	5	17.2	3414	14	O10383	2.03e+02	POLYPROTEIN.	2.03e+02
495	5	17.2	3414	14	O88489	2.03e+02	WUGSC_HDJ0751H13.1 PR	2.03e+02
496	5	17.2	4123	4	O75851	2.03e+02	FXBC.	2.03e+02
497	5	17.2	4976	2	O87314	2.03e+02	ZONADHESIN.	2.03e+02
498	5	17.2	5376	11	O88799	2.03e+02	NEBULIN.	2.03e+02
499	5	17.2	5669	4	O15346	2.03e+02	TITIN, HEART ISOFORM N	2.03e+02
500	5	17.2	5692	4	O10466	2.03e+02	MYOCHROME OXIDASE I (	2.03e+02
501	4	13.8	37	8	O33433	4.75e+03	MHC-DQA1*02 (FRAGMENT)	4.75e+03
502	4	13.8	62	7	O30503	4.75e+03	MHC-DQA1*03 (FRAGMENT)	4.75e+03
503	4	13.8	62	7	O30995	4.75e+03	MHC-DQA1*03 (FRAGMENT)	4.75e+03
504	4	13.8	62	2	O50163	4.75e+03	U296R.	4.75e+03
505	4	13.8	63	7	O30505	4.75e+03	MHC-DQA2*02 (FRAGMENT)	4.75e+03
506	4	13.8	63	7	O78024	4.75e+03	MHC CLASS II ANTIGEN (	4.75e+03
507	4	13.8	63	7	O30997	4.75e+03	MHC-DQA2*01 (FRAGMENT)	4.75e+03
508	4	13.8	63	7	O30574	4.75e+03	MHC-DQA1*01 (FRAGMENT)	4.75e+03
509	4	13.8	71	2	O68688	4.75e+03	PROBABLE SPORE GERMINA	4.75e+03
510	4	13.8	85	2	O89050	4.75e+03	ORF86.	4.75e+03
511	4	13.8	90	6	O62737	4.75e+03	EXTRACELLULAR RELATED	4.75e+03
512	4	13.8	90	6	O28835	4.75e+03	CGMP-GATED RETINAL PHO	4.75e+03
513	4	13.8	94	2	O25647	4.75e+03	HYPOTHETICAL 11.4 KD P	4.75e+03
514	4	13.8	96	2	O05352	4.75e+03	SIMILAR TO GROES PROTE	4.75e+03
515	4	13.8	97	10	O24572	4.75e+03	ESRA1 PROTEIN (FRAGMEN	4.75e+03
516	4	13.8	100	1	O26367	4.75e+03	RIBOSOMAL PROTEIN S24.	4.75e+03
517	4	13.8	100	8	O33558	4.75e+03	NH2 TERMINUS UNCERTAIN	4.75e+03
518	4	13.8	101	4	O14553	4.75e+03	R31240.1 (FRAGMENT).	4.75e+03
519	4	13.8	101	8	O35945	4.75e+03	YEAST (S.UVARUM) MITOC	4.75e+03
520	4	13.8	112	2	P71491	4.75e+03	HYPOTHETICAL PROTEIN (	4.75e+03
521	4	13						

532	4	13.8	139	2	P73322	YCF21.	4.75e+03	605	4	13.8	264	7	Q31162	MHC CLASS II E-BETA-F.	4.75e+03
533	4	13.8	142	2	Q50813	HYPOTHETICAL 15.0 KD P	4.75e+03	606	4	13.8	265	2	Q48639	BGLR.	4.75e+03
534	4	13.8	149	4	Q14292	FAS SOLUBLE PROTEIN.	4.75e+03	607	4	13.8	266	7	Q30166	MHC CLASS II HLA-DR-BE	4.75e+03
535	4	13.8	154	5	Q76857	BCL7-LIKE PROTEIN.	4.75e+03	608	4	13.8	266	7	Q30166	MHC CLASS II DOA2 (BOL	4.75e+03
536	4	13.8	155	2	Q25826	INVASION PROTEIN (INVA	4.75e+03	609	4	13.8	266	7	Q30126	MHC CLASS II HLA-DR-BE	4.75e+03
537	4	13.8	155	8	Q37035	NADH DEHYDROGENASE SUB	4.75e+03	610	4	13.8	266	7	Q31414	MHC CLASS II BETA CHAI	4.75e+03
538	4	13.8	155	2	P74426	HYPOTHETICAL 17.2 KD P	4.75e+03	611	4	13.8	266	7	Q30499	INTEGRAL MEMBRANE PROT	4.75e+03
539	4	13.8	160	2	Q84540	ACYL-COA THIOESTER HYD	4.75e+03	612	4	13.8	266	7	Q30835	MHC OVAR-DRB1 PRECURSO	4.75e+03
540	4	13.8	162	2	Q87116	ORF 4.	4.75e+03	613	4	13.8	266	7	Q30109	MHC CLASS II HLA-DR BE	4.75e+03
541	4	13.8	163	2	Q55178	INTEGRAL MEMBRANE 18.6 KD P	4.75e+03	614	4	13.8	266	7	Q30120	MHC CLASS II HLA-DR-BE	4.75e+03
542	4	13.8	169	7	Q30431	MEMBRANE LIPID SYNTHES	4.75e+03	615	4	13.8	268	5	Q26561	NF-YA SUBUNIT.	4.75e+03
543	4	13.8	169	2	Q46578	MEMBRANE LIPID SYNTHES	4.75e+03	616	4	13.8	268	5	Q86780	PUTATIVE LIPOPROTEIN.	4.75e+03
544	4	13.8	176	5	Q26633	E-ABL MRNA (FRAGMENT).	4.75e+03	617	4	13.8	269	5	Q18934	D1046.3 PROTEIN.	4.75e+03
545	4	13.8	178	5	P91101	SIMILAR TO MARINER TRA	4.75e+03	618	4	13.8	272	2	Q68527	HYPOTHETICAL 30.1 KD P	4.75e+03
546	4	13.8	180	2	Q59793	TRYPTOPHAN SYNTHASE AL	4.75e+03	619	4	13.8	274	7	P79556	MHC CLASS I HLA-A (FRA	4.75e+03
547	4	13.8	180	4	Q15107	PREGNANCY-SPECIFIC BET	4.75e+03	620	4	13.8	275	2	Q53442	STEAROYL-ACYL CARRIER	4.75e+03
548	4	13.8	183	2	O85605	PUTATIVE EXCINUCLEASE	4.75e+03	621	4	13.8	275	2	P74492	EXODOXRIBONUCLEASE I	4.75e+03
549	4	13.8	184	2	Q31333	BC541A PROTEIN (FRAGME	4.75e+03	622	4	13.8	275	2	Q69051	PTXA.	4.75e+03
550	4	13.8	186	5	Q19670	P21C3.6 PROTEIN.	4.75e+03	623	4	13.8	276	5	Q02249	F23B2.7 PROTEIN.	4.75e+03
551	4	13.8	189	4	Q15106	PREGNANCY-SPECIFIC BET	4.75e+03	624	4	13.8	276	2	P73137	ESTERASE.	4.75e+03
552	4	13.8	190	2	Q67509	GLIDING MOTILITY PROTE	4.75e+03	625	4	13.8	276	5	Q17808	COSMID C08A9.	4.75e+03
553	4	13.8	191	5	Q18708	PEPTIDYL-TRNA HYDROLAS	4.75e+03	626	4	13.8	276	5	Q43942	ACTIVATED PROTEIN KINA	4.75e+03
554	4	13.8	202	2	O83975	PEPTIDYL-TRNA HYDROLAS	4.75e+03	627	4	13.8	277	1	O58047	277RA LONG HYPOTHETICA	4.75e+03
555	4	13.8	203	5	Q27815	PHOSPHOENOLPYRUVATE CA	4.75e+03	628	4	13.8	280	5	Q27030	32 KDA SURFACE ANTIGEN	4.75e+03
556	4	13.8	205	2	Q44152	MEMBRANE-BOUND ALCOHOL	4.75e+03	629	4	13.8	282	2	Q45815	NAD-DEPENDENT BETA-HYD	4.75e+03
557	4	13.8	205	2	Q34381	TRANSCRIPTIONAL REGULA	4.75e+03	630	4	13.8	283	2	P73096	HYPOTHETICAL 31.9 KD P	4.75e+03
558	4	13.8	206	2	P75433	TYPE 1 RESTRICTION ENZ	4.75e+03	631	4	13.8	286	10	Q53404	HYPOHETICAL 31.9 KD P	4.75e+03
559	4	13.8	209	2	P76557	FROM BASES 3180257 TO	4.75e+03	632	4	13.8	286	2	O65597	HYPOHETICAL 31.9 KD P	4.75e+03
560	4	13.8	210	8	O47536	CYTOCHROME OXIDASE I (	4.75e+03	633	4	13.8	288	2	O85489	NCOI RESTRICTION ENDON	4.75e+03
561	4	13.8	211	8	O47527	CYTOCHROME OXIDASE I (	4.75e+03	634	4	13.8	289	2	O25390	LIPOPOLYSACCHARIDE BIO	4.75e+03
562	4	13.8	213	10	Q49929	FORE PROTEIN OF 24 KD	4.75e+03	635	4	13.8	293	5	O16236	F07G11.8 PROTEIN.	4.75e+03
563	4	13.8	214	8	O47532	CYTOCHROME OXIDASE I (	4.75e+03	636	4	13.8	299	5	Q21490	C. ELEGANS DNA-BINDING	4.75e+03
564	4	13.8	217	8	O47525	CYTOCHROME OXIDASE I (	4.75e+03	637	4	13.8	304	5	Q22480	COSMID T14B4.	4.75e+03
565	4	13.8	219	3	Q02172	120 KD GLYCOPROTEIN AN	4.75e+03	638	4	13.8	306	2	O69358	2,3-DIHYDROXYBIPHENYL	4.75e+03
566	4	13.8	220	7	P79467	MHC CLASS II DR BETA-C	4.75e+03	639	4	13.8	308	4	O15071	OXYTOCIN RECEPTOR (FRA	4.75e+03
567	4	13.8	221	2	O85559	HYPOTHETICAL 24.5 KD P	4.75e+03	640	4	13.8	308	5	O44871	R06B10.3 PROTEIN (FRAG	4.75e+03
568	4	13.8	222	2	O52733	HYPOTHETICAL 24.9 KD P	4.75e+03	641	4	13.8	309	5	O45561	F54B8.3 PROTEIN.	4.75e+03
569	4	13.8	223	2	O52255	FLHA (FRAGMENT).	4.75e+03	642	4	13.8	309	2	O50643	SALR.	4.75e+03
570	4	13.8	225	2	Q31235	OMLA (FRAGMENT).	4.75e+03	643	4	13.8	310	5	O16367	F29G9.4 PROTEIN.	4.75e+03
571	4	13.8	225	10	O64772	T1F9.3.	4.75e+03	644	4	13.8	313	5	Q18054	SIMILARITY TO BASIC HE	4.75e+03
572	4	13.8	225	5	P90525	SECG (FRAGMENT).	4.75e+03	645	4	13.8	315	2	O07227	GAMMA-HEMOLYSIN H-GAMM	4.75e+03
573	4	13.8	226	2	P90384	HYPOTHETICAL 24.1 KD P	4.75e+03	646	4	13.8	316	2	Q52464	ALGINATE SYNTHESIS-REL	4.75e+03
574	4	13.8	226	10	O65377	FL2F1.11.	4.75e+03	647	4	13.8	317	2	Q59244	L-LACTATE DEHYDROGENAS	4.75e+03
575	4	13.8	227	4	Q43970	ARAC-LIKE PROTEIN.	4.75e+03	648	4	13.8	318	2	O54351	OUTER MEMBRANE PROTEIN	4.75e+03
576	4	13.8	227	7	Q30065	MHC CLASS II HLA-DQ-AL	4.75e+03	649	4	13.8	319	2	O05804	RNA-DIRECTED DNA POLYM	4.75e+03
577	4	13.8	230	2	Q49630	HYPOTHETICAL 24.4 KD P	4.75e+03	650	4	13.8	321	2	P66521	ORF4.	4.75e+03
578	4	13.8	231	2	O46471	HYDROGENASE CYTOCHROME	4.75e+03	651	4	13.8	324	2	O31392	GLUCOSE KINASE (EC 2.7	4.75e+03
579	4	13.8	232	7	Q30086	MHC CLASS II HLA-DQ-AL	4.75e+03	652	4	13.8	324	8	O63278	CYTOCHROME OXIDASE SUB	4.75e+03
580	4	13.8	233	2	O34954	YOLB.	4.75e+03	653	4	13.8	324	8	O63277	CYTOCHROME OXIDASE SUB	4.75e+03
581	4	13.8	233	2	O25334	ABC TRANSPORTER, ATP-B	4.75e+03	654	4	13.8	327	2	O69638	PUTATIVE EPOXIDE HYDRO	4.75e+03
582	4	13.8	233	3	Q12483	HYPOTHETICAL 27.0 KD P	4.75e+03	655	4	13.8	330	2	Q55280	PERIPLASMIC-BINDING PR	4.75e+03
583	4	13.8	233	2	Q04138	REPLICATION PROTEIN A.	4.75e+03	656	4	13.8	330	2	O44299	OUTER MEMBRANE PROTEIN	4.75e+03
584	4	13.8	235	2	O45021	REPEATED DNA ELEMENT,	4.75e+03	657	4	13.8	331	5	O26469	HOMEOIC PROTEIN.	4.75e+03
585	4	13.8	237	7	Q27992	MHC CLASS II HLA-DR2,	4.75e+03	658	4	13.8	334	4	O60386	ZN-ALPHA2-GLYCOPROTEIN	4.75e+03
586	4	13.8	237	7	O19507	MHC CLASS II HLA-DR2,	4.75e+03	659	4	13.8	336	2	P73160	CYTOCHROME OXIDASE D S	4.75e+03
587	4	13.8	238	6	Q29020	APOLIPOPROTEIN B GENE	4.75e+03	660	4	13.8	339	5	P91084	COSMID C23H3.	4.75e+03
588	4	13.8	240	4	Q13178	PREGNANCY-SPECIFIC GLY	4.75e+03	661	4	13.8	341	5	Q32169	YUSC PROTEIN.	4.75e+03
589	4	13.8	243	2	O68406	POLYSIALIC ACID TRANSP	4.75e+03	662	4	13.8	341	5	O17770	F02E9.8 PROTEIN.	4.75e+03
590	4	13.8	245	2	O54231	UNIDENTIFIED ORF (FRAG	4.75e+03	663	4	13.8	343	7	O46826	MHC CLASS I ANTIGEN.	4.75e+03
591	4	13.8	248	2	O25785	GLUTAMINE ABC TRANSPOR	4.75e+03	664	4	13.8	344	5	O16378	M03F8.1 PROTEIN.	4.75e+03
592	4	13.8	251	4	O75797	ORF-2.	4.75e+03	665	4	13.8	345	5	O76465	MEROZOITE CAPPING PROT	4.75e+03
593	4	13.8	251	2	O66373	PHOTOSYNTHETIC REACTIO	4.75e+03	666	4	13.8	346	8	O19859	RIBULOSE BIPHOSPHATE	4.75e+03
594	4	13.8	253	4	O15157	DOLICHO MONOPHOSPHATE	4.75e+03	667	4	13.8	348	5	Q21558	M18.6 PROTEIN.	4.75e+03
595	4	13.8	254	8	O33935	CYTOCHROME OXIDASE I (	4.75e+03	668	4	13.8	351	5	P90892	COSMID F56E3.	4.75e+03
596	4	13.8	254	8	Q33934	CYTOCHROME OXIDASE I (	4.75e+03	669	4	13.8	351	8	O79033	CYTOCHROME OXIDASE I (	4.75e+03
597	4	13.8	255	7	Q30308	MHC CLASS II DOA2 PREC	4.75e+03	670	4	13.8	352	2	O68498	ADAPTIVE RESPONSE REGU	4.75e+03
598	4	13.8	257	2	O67644	HYPOTHETICAL 29.5 KD P	4.75e+03	671	4	13.8	353	8	O47467	CYTOCHROME OXIDASE SUB	4.75e+03
599	4	13.8	257	7	Q30344	INTEGRAL MEMBRANE GLYC	4.75e+03	672	4	13.8	353	8	O47469	CYTOCHROME OXIDASE SUB	4.75e+03
600	4	13.8	257	7	Q30343	INTEGRAL MEMBRANE GLYC	4.75e+03	673	4	13.8	353	8	O47464	CYTOCHROME OXIDASE SUB	4.75e+03
601	4	13.8	257	5	O20426	COSMID F45E1.	4.75e+03	674	4	13.8	354	2	O52539	MEMBRANE PROTEIN.	4.75e+03
602	4	13.8	261	7	O28668	MHC CLASS II HISTOCOMP	4.75e+03	675	4	13.8	354	7	Q95526	CLASS I HISTOCOMPATIBI	4.75e+03
603	4	13.8	261	7	Q31112	HISTOCOMPATIBILITY 2,	4.75e+03	676	4	13.8	354	7	Q95525	CLASS I HISTOCOMPATIBI	4.75e+03
604	4	13.8	261	2	O53774	HYPOTHETICAL 27.3 KD P	4.75e+03	677	4	13.8	354	2	Q56844	YSCU.	4.75e+03





824	4	13.8	476	8	046962	RIBULOSE BIPHOSPHATE	4.75e+03	897	4	13.8	558	2	P74702	HEAT SHOCK PROTEIN.	4.75e+03
825	4	13.8	476	8	046963	RIBULOSE BIPHOSPHATE	4.75e+03	898	4	13.8	567	5	Q18322	COSMID C30B5.	4.75e+03
826	4	13.8	477	8	096000	CYTOCHROME OXIDASE (EC	4.75e+03	899	4	13.8	571	5	Q07431	RETROPOSON (EC 2.7.7.4	4.75e+03
827	4	13.8	478	8	047021	RIBULOSE BIPHOSPHATE	4.75e+03	900	4	13.8	578	4	Q44326	HMP-2.	4.75e+03
828	4	13.8	480	8	032338	RIBULOSE BIPHOSPHATE	4.75e+03	901	4	13.8	582	2	Q48674	NISP (NISP PROTEIN).	4.75e+03
829	4	13.8	481	2	087729	TRKH PROTEIN.	4.75e+03	902	4	13.8	584	4	P78424	RETINA-DERIVED POU-DOM	4.75e+03
830	4	13.8	482	3	Q05911	HYPOTHETICAL 54.5 KD P	4.75e+03	903	4	13.8	585	5	Q44402	F37CA.3 PROTEIN.	4.75e+03
831	4	13.8	482	4	Q51149	KIAA0660 PROTEIN.	4.75e+03	904	4	13.8	586	5	Q07492	PROTEIN ESAG 10.	4.75e+03
832	4	13.8	483	5	P91568	SIMILAR TO PROTEIN-TYR	4.75e+03	905	4	13.8	595	2	Q85672	GLUCAMYLASE (EC 3.2.1	4.75e+03
833	4	13.8	487	2	Q50901	BETA-1,4-GLYCANASE.	4.75e+03	906	4	13.8	598	5	Q44447	CO2B10.5 PROTEIN.	4.75e+03
834	4	13.8	487	2	Q50304	OUTER MEMBRANE PROTEIN	4.75e+03	907	4	13.8	698	4	Q16434	NA+/H+ EXCHANGER.	4.75e+03
835	4	13.8	488	2	085775	PUTATIVE REGULATORY PR	4.75e+03	908	4	13.8	700	4	Q13269	CAMP RESPONSIVE ELEMEN	4.75e+03
836	4	13.8	490	2	084234	SODIUM-DEPENDENT AMINO	4.75e+03	909	4	13.8	700	5	Q25884	INTERSPERSED REPEAT AN	4.75e+03
837	4	13.8	491	5	Q45988	ZK1037.6 PROTEIN.	4.75e+03	910	4	13.8	703	2	Q33376	TRANSFERRIN-BINDING PR	4.75e+03
838	4	13.8	493	5	077017	PUTATIVE AMYLASE-RELAT	4.75e+03	911	4	13.8	705	2	Q31988	SUBLANCIN 168 LANTIBIO	4.75e+03
839	4	13.8	495	4	Q43814	SERUM RESPONSE FACTOR-	4.75e+03	912	4	13.8	707	2	Q30671	SUBLANCIN TRANSPORTER	4.75e+03
840	4	13.8	497	4	075511	SUPPRESSIN.	4.75e+03	913	4	13.8	708	3	Q12027	CHROMOSOME IV READING PR	4.75e+03
841	4	13.8	501	10	Q22815	PUTATIVE MLO PROTEIN.	4.75e+03	914	4	13.8	708	2	Q50930	TRANSFERRIN BINDING PR	4.75e+03
842	4	13.8	502	5	Q45485	F38A1.5 PROTEIN.	4.75e+03	915	4	13.8	710	5	P91272	CODED FOR BY C. ELEGAN	4.75e+03
843	4	13.8	506	5	Q17381	CE-FKH-1.	4.75e+03	916	4	13.8	715	2	Q83248	SIGMA FACTOR S1GG REGU	4.75e+03
844	4	13.8	507	1	Q26749	CONSERVED PROTEIN.	4.75e+03	917	4	13.8	719	2	Q48868	ABC EXPORTER.	4.75e+03
845	4	13.8	508	2	P73375	L-THREONINE DEAMINASE.	4.75e+03	918	4	13.8	722	2	P96091	SERINE PROTEASE (DENTI	4.75e+03
846	4	13.8	508	10	Q49714	HYPOTHETICAL 55.2 KD P	4.75e+03	919	4	13.8	726	4	Q15152	PLAKOPHILIN.	4.75e+03
847	4	13.8	511	1	Q29598	TYPE II SECRETION SYST	4.75e+03	920	4	13.8	729	5	Q18660	BZIP TRANSCRIPTION FAC	4.75e+03
848	4	13.8	522	5	Q18187	W09G3.3 PROTEIN.	4.75e+03	921	4	13.8	731	2	Q86557	PUTATIVE ATP /GTP BIND	4.75e+03
849	4	13.8	523	5	Q26769	INVARIANT SURFACE GLYC	4.75e+03	922	4	13.8	732	4	Q13032	GABA/NORADRENALINE TRA	4.75e+03
850	4	13.8	525	5	Q02101	SIMILARITY TO MOUSE D1	4.75e+03	923	4	13.8	732	5	Q17619	C29F7.6 PROTEIN.	4.75e+03
851	4	13.8	536	10	Q65837	LYCOPENE EPSILON-CYCLA	4.75e+03	924	4	13.8	739	5	P91060	SIMILAR TO PEROXIDASE	4.75e+03
852	4	13.8	527	2	Q88013	HYPOTHETICAL 56.9 KD P	4.75e+03	925	4	13.8	740	5	Q01763	SIMILARITY TOTHE HUMAN	4.75e+03
853	4	13.8	530	5	Q18303	ZK849.4 PROTEIN.	4.75e+03	926	4	13.8	741	2	Q24875	ATP-DEPENDENT C1P PROT	4.75e+03
854	4	13.8	534	2	Q34674	YTGP.	4.75e+03	927	4	13.8	745	5	Q61458	MEDEA.	4.75e+03
855	4	13.8	545	5	Q27618	CARBOXYLESTERASE (EC 3	4.75e+03	928	4	13.8	752	5	Q94814	SPOROZOITE SURFACE PRO	4.75e+03
856	4	13.8	545	5	Q27616	CARBOXYLESTERASE (EC 3	4.75e+03	929	4	13.8	753	4	Q75175	KIAA0691 PROTEIN.	4.75e+03
857	4	13.8	545	5	Q27623	CARBOXYLESTERASE (EC 3	4.75e+03	930	4	13.8	755	5	Q17612	C29F3.1 PROTEIN.	4.75e+03
858	4	13.8	545	5	Q27625	CARBOXYLESTERASE (EC 3	4.75e+03	931	4	13.8	756	4	Q75154	KIAA0665 PROTEIN.	4.75e+03
859	4	13.8	545	5	Q16171	CARBOXYLESTERASE-SC.	4.75e+03	932	4	13.8	759	2	Q05143	ENDOSULCANASE A PRECUR	4.75e+03
860	4	13.8	546	1	Q28405	HYPOTHETICAL 62.8 KD P	4.75e+03	933	4	13.8	762	6	Q46406	COPPER AMINE OXIDASE P	4.75e+03
861	4	13.8	547	5	Q45965	Y61B9A.1 PROTEIN.	4.75e+03	934	4	13.8	767	5	Q93325	C33G3.1 PROTEIN.	4.75e+03
862	4	13.8	550	5	Q19680	F21D5.1 PROTEIN.	4.75e+03	935	4	13.8	774	2	Q32215	YVGS PROTEIN.	4.75e+03
863	4	13.8	558	5	Q26631	ENDO16 (FRAGMENT).	4.75e+03	936	4	13.8	778	2	P95149	HYPOTHETICAL 83.2 KD P	4.75e+03
864	4	13.8	559	2	Q51513	PHA-SYNTHASE1.	4.75e+03	937	4	13.8	790	5	Q18321	SIMILAR TO ADRENERGIC	4.75e+03
865	4	13.8	553	2	Q47519	PYRUVATE OXIDASE.	4.75e+03	938	4	13.8	791	4	Q75120	KIAA0625 PROTEIN (FRAG	4.75e+03
866	4	13.8	553	2	P71716	META.	4.75e+03	939	4	13.8	793	2	Q48571	X-PROLYL DIPEPTIDYL AM	4.75e+03
867	4	13.8	566	2	P71093	NEOULLULANASE.	4.75e+03	940	4	13.8	796	5	Q20651	CODED FOR BY C. ELEGAN	4.75e+03
868	4	13.8	572	2	Q47514	PYRUVATE OXIDASE.	4.75e+03	941	4	13.8	799	5	Q27591	INTEGRIN BETA SUBUNIT	4.75e+03
869	4	13.8	578	2	Q25807	MULTIDRUG RESISTANCE P	4.75e+03	942	4	13.8	819	4	Q75366	ADVILLIN.	4.75e+03
870	4	13.8	579	5	Q94049	T13F2.6 PROTEIN.	4.75e+03	943	4	13.8	824	5	Q17763	F01G10.5 PROTEIN.	4.75e+03
871	4	13.8	582	10	Q65670	PUTATIVE L-ASCORBATE O	4.75e+03	944	4	13.8	839	2	Q55023	CLPC.	4.75e+03
872	4	13.8	584	2	Q68318	RESPONSE REGULATOR.	4.75e+03	945	4	13.8	843	5	P91298	COSMID F36H5.	4.75e+03
873	4	13.8	585	2	Q07550	HYPOTHETICAL 65.1 KD P	4.75e+03	946	4	13.8	857	5	Q22878	COSMID B0507.	4.75e+03
874	4	13.8	592	10	Q49006	PUTATIVE PECTIN METHYL	4.75e+03	947	4	13.8	869	2	Q47912	1,4-B-D-GLUCAN GLUCOXY	4.75e+03
875	4	13.8	593	4	Q00499	AMPHIPHYSIN II.	4.75e+03	948	4	13.8	870	4	Q60309	KIAA0563 PROTEIN.	4.75e+03
876	4	13.8	593	2	P75706	FROM BASES 464774 TO 4	4.75e+03	949	4	13.8	873	8	Q37370	CYTOCHROME C OXIDASE P	4.75e+03
877	4	13.8	595	2	Q84799	DNA PRIMAISE.	4.75e+03	950	4	13.8	885	3	Q74678	CHITIN SYNTHASE 3.	4.75e+03
878	4	13.8	595	2	Q31825	YNGE PROTEIN.	4.75e+03	951	4	13.8	900	3	Q74925	PUTATIVE VACUOLAR MEMB	4.75e+03
879	4	13.8	596	5	Q62306	NA/CA,K-EXCHANGER.	4.75e+03	952	4	13.8	902	10	Q64736	T27E13.16 PROTEIN.	4.75e+03
880	4	13.8	605	2	Q84668	DNA GYRASE SUBUNIT B.	4.75e+03	953	4	13.8	913	3	Q12151	HYPOTHETICAL 100.3 KD	4.75e+03
881	4	13.8	607	5	Q19321	RIIAR.4 PROTEIN.	4.75e+03	954	4	13.8	924	2	Q84754	ALANYL TRNA SYNTHETASE	4.75e+03
882	4	13.8	608	1	Q28171	SIGNAL-TRANSDUCING HIS	4.75e+03	955	4	13.8	940	4	Q00405	FB19 PROTEIN.	4.75e+03
883	4	13.8	610	5	Q76486	GLUCOSE TRANSPORTER.	4.75e+03	956	4	13.8	959	3	Q74290	COPALYL DIPHOSPHATE SY	4.75e+03
884	4	13.8	618	2	Q44881	P66 PROTEIN PRECURSOR.	4.75e+03	957	4	13.8	1008	5	Q77391	MAL3P6.4 PROTEIN.	4.75e+03
885	4	13.8	624	5	Q61393	GLYC5 (EC 2.4.1.41) (P	4.75e+03	958	4	13.8	1021	5	Q15733	HYPOTHETICAL 113.3 KD	4.75e+03
886	4	13.8	626	5	Q61392	GLY5B (EC 2.4.1.41) (P	4.75e+03	959	4	13.8	1036	4	Q75119	KIAA0623 PROTEIN.	4.75e+03
887	4	13.8	633	5	Q61104	HYPOTHETICAL 70.0 KD P	4.75e+03	960	4	13.8	1041	5	Q01351	REVERSE TRANSCRIPTASE	4.75e+03
888	4	13.8	636	3	Q13858	HYPOTHETICAL 69.8 KD P	4.75e+03	961	4	13.8	1045	5	Q20047	SIMILAR TO S. CEREVISI	4.75e+03
889	4	13.8	638	2	Q51565	NORD PROTEIN.	4.75e+03	962	4	13.8	1051	2	P70745	GLUCODEXTRANASE PRECUR	4.75e+03
890	4	13.8	645	2	Q51869	HEAT SHOCK PROTEIN DNA	4.75e+03	963	4	13.8	1075	5	Q19540	COSMID F18A1.	4.75e+03
891	4	13.8	646	10	Q64877	F411.24 PROTEIN.	4.75e+03	964	4	13.8	1081	4	Q60303	KIAA0556 PROTEIN (FRAG	4.75e+03
892	4	13.8	650	5	Q21087	K01C8.3 PROTEIN.	4.75e+03	965	4	13.8	1082	3	Q13398	P-TYPE ATPASE 2.	4.75e+03
893	4	13.8	650	2	Q54022	INVASION PROTEIN (FRAG	4.75e+03	966	4	13.8	1087	5	Q22490	SIMILAR TO DOPAMINE RE	4.75e+03
894	4	13.8	652	5	Q01435	SIMILAR TO PROTEIN-TYR	4.75e+03	967	4	13.8	1093	4	Q00203	AP-3 COMPLEX BETA3A SU	4.75e+03
895	4	13.8	652	3	Q47974	HYPOTHETICAL 75.7 KD P	4.75e+03	968	4	13.8	1104	4	Q60460	MUCIN (FRAGMENT).	4.75e+03
896	4	13.8	653	5	Q22553	SIMILAR TO SERINE.THRE	4.75e+03	969	4	13.8	1121	5	Q02252	F23B2.11 PROTEIN.	4.75e+03

970 4 13.8 1124 4 Q13800 4.75e+03  
971 4 13.8 1131 5 Q18529 4.75e+03  
972 4 13.8 1156 6 Q46371 4.75e+03  
973 4 13.8 1186 5 Q17666 4.75e+03  
974 4 13.8 1188 5 Q62302 4.75e+03  
975 4 13.8 1194 4 Q99790 4.75e+03  
976 4 13.8 1248 4 Q60610 4.75e+03  
977 4 13.8 1254 5 Q26599 4.75e+03  
978 4 13.8 1256 4 Q75085 4.75e+03  
979 4 13.8 1261 2 Q54043 4.75e+03  
980 4 13.8 1287 3 Q06047 4.75e+03  
981 4 13.8 1461 4 Q00340 4.75e+03  
982 4 13.8 1502 3 Q06574 4.75e+03  
983 4 13.8 1502 3 Q06574 4.75e+03  
984 4 13.8 1521 5 Q93148 4.75e+03  
985 4 13.8 1538 2 Q53395 4.75e+03  
986 4 13.8 1538 2 Q53234 4.75e+03  
987 4 13.8 1613 4 Q75581 4.75e+03  
988 4 13.8 1698 5 Q94438 4.75e+03  
989 4 13.8 1706 2 Q51839 4.75e+03  
990 4 13.8 1807 5 Q17428 4.75e+03  
991 4 13.8 2018 5 Q20487 4.75e+03  
992 4 13.8 2235 5 Q61077 4.75e+03  
993 4 13.8 2233 5 Q94711 4.75e+03  
994 4 13.8 2437 5 Q77393 4.75e+03  
995 4 13.8 2893 2 Q25063 4.75e+03  
996 4 13.8 3351 5 Q94907 4.75e+03  
997 4 13.8 3722 2 Q94873 4.75e+03  
998 4 13.8 3772 2 Q73139 4.75e+03  
999 4 13.8 4151 2 Q53490 4.75e+03  
1000 4 13.8 4641 4 Q75592 4.75e+03

## ALIGNMENTS

RESULT 1  
ID O64696 PRELIMINARY; PRT; 401 AA.  
AC O64696;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
DE T31E10.15 PROTEIN.  
GN T31E10.15.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA ROUNSLEY S.D., KAUL S., LIN X., KETCHUM K.A., CROSBY M.L.,  
RA BRANDON R.C., SKES S.M., MASON T.M., KERLAGE A.R., ADAMS M.D.,  
RA SOMERVILLE C.R., VENTER J.C.;  
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
DR EMBL; AC004077; G3128209;  
SQ SEQUENCE 401 AA; 43943 MW; 41E6DF15 CRC32;

Query Match 24.1%; Score 7; DB 10; Length 401;  
Best Local Similarity 100.0%; Pred. No. 7.57e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 SQKOGG 86  
|||||  
Qy 15 SQKOGG 21

RESULT 2  
ID O44092 PRELIMINARY; PRT; 90 AA.  
AC O44092;

DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE VACUOLAR ATPASE 14KD SUBUNIT (FRAGMENT).

GN VHA14.  
OS DROSOPHILA SUBOBSCURA (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ZENG L.-W., COMERON J.M., CHEN B., KREITMAN M.;  
RL GENETICA 0:0-0(1997).  
DR EMBL; AF025797; G2655145;  
FT NON\_TER 1  
FT NON\_TER 90  
SQ SEQUENCE 90 AA; 10128 MW; F32B0EB5 CRC32;

Query Match 20.7%; Score 6; DB 5; Length 90;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 VVDKNT 31  
|||||  
Qy 23 VVDKNT 28

RESULT 3  
ID O44091 PRELIMINARY; PRT; 90 AA.  
AC O44091;

DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE VACUOLAR ATPASE 14KD SUBUNIT (FRAGMENT).

GN VHA14.  
OS DROSOPHILA PSEUDOSCURA (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ZENG L.-W., COMERON J.M., CHEN B., KREITMAN M.;  
RL GENETICA 0:0-0(1997).  
DR EMBL; AF025796; G2655143;  
FT NON\_TER 1  
FT NON\_TER 90  
SQ SEQUENCE 90 AA; 10128 MW; F32B0EB5 CRC32;

Query Match 20.7%; Score 6; DB 5; Length 90;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 VVDKNT 31  
|||||  
Qy 23 VVDKNT 28

RESULT 4  
ID O85708 PRELIMINARY; PRT; 111 AA.  
AC O85708;

DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE THLC (FRAGMENT).

GN THLC.  
OS CLOSTRIDIUM ACETOBUTYLICUM.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;  
OC CLOSTRIDIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 792;  
RX MEDLINE; 96025601.

RA DURR P., FISCHER R.J., KUHN A., LORENZ K., SCHREIBER W.,  
RA STURZENHOEFER B., ULLMANN S., WINZER K., SAUER U.;  
RT "Solventogenic enzymes of Clostridium acetobutylicum: catalytic  
properties, genetic organization, and transcriptional regulation."  
RL FEMS MICROBIOL. REV. 17:251-262(1995).

```

RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-DSM 792;
RA  WINZER K., LORENZ K., DURRE P.;
RT  "Cloning, sequencing, and molecular analysis of two thiolase genes
RL  from Clostridium acetobutylicum DSM 792";
RL  SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; AF072735; G3309207; -.
FT  NON_TER 111 111
SQ  SEQUENCE 111 AA; 12966 MW; CFIF99AE CRC32;

Query Match      20.7%; Score 6; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  98 VDKNTK 103
QY  24 VDKNTK 29
    |||||

RESULT      5
ID  Q65439      PRELIMINARY;      PRT;      173 AA.
AC  Q65439;
DT  01-NOV-1996 (TREMREL. 01, CREATED)
DT  01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT  01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE  ORF L2.
OS  BEET CURLY TOP VIRUS (BCTV).
OC  VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; CURTOVIRUS.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CFH;
RX  MEDLINE; 94220740.
RA  STENGER D.C.;
RT  "Complete nucleotide sequence of the hypervirulent CFH strain of beet
RL  curly top virus.";
RL  MOL. PLANT MICROBE INTERACT. 7:154-157(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CFH;
RA  STENGER D.C.;
RL  SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; U02211; G522157; -.
SQ  SEQUENCE 173 AA; 19545 MW; 34C3AF74 CRC32;

Query Match      20.7%; Score 6; DB 14; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  132 QPPQIL 137
QY  5 QPPQIL 10
    |||||

RESULT      6
ID  Q72691      PRELIMINARY;      PRT;      173 AA.
AC  Q72691;
DT  01-AUG-1998 (TREMREL. 07, CREATED)
DT  01-AUG-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
DT  01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE  C2 PROTEIN.
OS  BEET CURLY TOP VIRUS (BCTV).
OC  VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; CURTOVIRUS.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE-LEAF;
RA  BRIDDON R.W.;
RL  SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; X97203; E236491; -.
SQ  SEQUENCE 173 AA; 19518 MW; 2EBE0A18 CRC32;

Query Match      20.7%; Score 6; DB 14; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-DSM 792;
RA  WINZER K., LORENZ K., DURRE P.;
RT  "Cloning, sequencing, and molecular analysis of two thiolase genes
RL  from Clostridium acetobutylicum DSM 792";
RL  SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; AF072735; G3309207; -.
FT  NON_TER 111 111
SQ  SEQUENCE 111 AA; 12966 MW; CFIF99AE CRC32;

Query Match      20.7%; Score 6; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  98 VDKNTK 103
QY  24 VDKNTK 29
    |||||

RESULT      5
ID  Q65439      PRELIMINARY;      PRT;      173 AA.
AC  Q65439;
DT  01-NOV-1996 (TREMREL. 01, CREATED)
DT  01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT  01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE  ORF L2.
OS  BEET CURLY TOP VIRUS (BCTV).
OC  VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; CURTOVIRUS.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CFH;
RX  MEDLINE; 94220740.
RA  STENGER D.C.;
RT  "Complete nucleotide sequence of the hypervirulent CFH strain of beet
RL  curly top virus.";
RL  MOL. PLANT MICROBE INTERACT. 7:154-157(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CFH;
RA  STENGER D.C.;
RL  SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; U02211; G522157; -.
SQ  SEQUENCE 173 AA; 19545 MW; 34C3AF74 CRC32;

Query Match      20.7%; Score 6; DB 14; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  132 QPPQIL 137
QY  5 QPPQIL 10
    |||||

RESULT      6
ID  Q72691      PRELIMINARY;      PRT;      173 AA.
AC  Q72691;
DT  01-AUG-1998 (TREMREL. 07, CREATED)
DT  01-AUG-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
DT  01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE  C2 PROTEIN.
OS  BEET CURLY TOP VIRUS (BCTV).
OC  VIRUSES; SSDNA VIRUSES; GEMINIVIRIDAE; CURTOVIRUS.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE-LEAF;
RA  BRIDDON R.W.;
RL  SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; X97203; E236491; -.
SQ  SEQUENCE 173 AA; 19518 MW; 2EBE0A18 CRC32;

Query Match      20.7%; Score 6; DB 14; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  132 QPPQIL 137
QY  5 QPPQIL 10
    |||||

RESULT      7
ID  O13619      PRELIMINARY;      PRT;      181 AA.
AC  O13619;
DT  01-JAN-1998 (TREMREL. 05, CREATED)
DT  01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT  01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE  P1028 PROTEIN.
GN  P1028.
OS  SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC  EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC  SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC  SCHIZOSACCHAROMYCES.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-972 H-;
RA  KUSHIDA N., YAMAZAKI S., TANAKA T., JINNO K., HAIKAWA Y., YAMAZAKI J.,
RA  YAMAMOTO S., SEKINE M., OGUCHI A., NAGAI Y., SAKAI M., AOKI K.,
RA  OGURA K., OTSUKA R., KUDOH Y., YANAGIDA M., MACHIDA M., ZHANG M.Q.;
RL  SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR  EMBL; AB004535; D102253; -.
SQ  SEQUENCE 181 AA; 20293 MW; C4A37153 CRC32;

Query Match      20.7%; Score 6; DB 3; Length 181;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  114 LLWDPS 119
QY  10 LLWDPS 15
    |||||

RESULT      8
ID  O14224      PRELIMINARY;      PRT;      205 AA.
AC  O14224;
DT  01-JUN-1998 (TREMREL. 06, CREATED)
DT  01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)
DT  01-JUN-1998 (TREMREL. 06, LAST ANNOTATION UPDATE)
DE  PUTATIVE RHO GDP-DISSOCIATION INHIBITOR (RHO GDI).
GN  SPAC6F12.06.
OS  SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC  EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC  SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC  SCHIZOSACCHAROMYCES.
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-972;
RA  BADCOCK K., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.;
RL  SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC  -!- FUNCTION: REGULATES THE GDP/GTP EXCHANGE REACTION OF THE RHO
CC  PROTEINS BY INHIBITING THE DISSOCIATION OF GDP FROM THEM, AND THE
CC  SUBSEQUENT BINDING OF GTP TO THEM (BY SIMILARITY).
CC  -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC  -!- SIMILARITY: TO OTHER EUKARYOTIC RHO GDP-DISSOCIATION INHIBITORS.
DR  EMBL; Z98533; E334111; -.
KW  HYPOTHETICAL PROTEIN; GTPASE ACTIVATION.
SQ  SEQUENCE 205 AA; 23032 MW; AA4B5FC3 CRC32;

Query Match      20.7%; Score 6; DB 3; Length 205;
Best Local Similarity 100.0%; Pred. No. 4.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db  138 GFVVDK 143
QY  21 GFVVDK 26
    |||||
```

RESULT 9  
ID P87309 PRELIMINARY; PRT; 207 AA.  
AC P87309;  
DT 01-JUN-1998 (TREMUREL. 06, CREATED)  
DT 01-JUN-1998 (TREMUREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMUREL. 06, LAST SEQUENCE UPDATE)  
DE HYPOTHETICAL 23.2 KD PROTEIN C31F10.08 IN CHROMOSOME II.  
GN SPC31F10.08.  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;  
OC SCHIZOSACCHAROMYCETES; SCHIZOSACCHAROMYCETACEAE;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA WOOD V., RAJANDREAM M.A., BARRELL B.G., POHL T.;  
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.  
DR EMBL; Z97204; E325256; -;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 207 AA; 23167 MW; A3176318 CRC32;

Query Match 20.7%; Score 6; DB 3; Length 207;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 64 VDKNTK 69  
|  
|  
|  
|  
|  
Qy 24 VDKNTK 29

RESULT 10  
ID O35297 PRELIMINARY; PRT; 295 AA.  
AC O35297;  
DT 01-JAN-1998 (TREMUREL. 05, CREATED)  
DT 01-JAN-1998 (TREMUREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMUREL. 07, LAST SEQUENCE UPDATE)  
DE SECRETED APOPTOSIS RELATED PROTEIN 1.  
GN SARPL.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MELKONIAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,  
RA FITZPATRICK P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;  
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.  
DR EMBL; AF017989; G2415421; -;  
SQ SEQUENCE 295 AA; 33483 MW; 0E4C8A8E CRC32;

Query Match 20.7%; Score 6; DB 11; Length 295;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 GOKQGG 266  
|  
|  
|  
|  
|  
Qy 16 GOKQGG 21

RESULT 11  
ID O08862 PRELIMINARY; PRT; 295 AA.  
AC O08862;  
DT 01-JUL-1997 (TREMUREL. 04, CREATED)  
DT 01-JUL-1997 (TREMUREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMUREL. 08, LAST SEQUENCE UPDATE)  
DE SECRETED FRIZZLED RELATED PROTEIN SRP-2.  
GN SRP2.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97250455.  
R# RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J., COPELAND N.G.,

RA JENKINS N.A., NATHANS J.;  
RT "A family of secreted proteins contains homology to the cysteine-rich  
RT ligand-binding domain of frizzled receptors";  
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).  
DR EMBL; U88567; G1946343; -;  
SQ SEQUENCE 295 AA; 33469 MW; 236B0DDA CRC32;

Query Match 20.7%; Score 6; DB 11; Length 295;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 GOKQGG 266  
|  
|  
|  
|  
|  
Qy 16 GOKQGG 21

RESULT 12  
ID P97299 PRELIMINARY; PRT; 295 AA.  
AC P97299;  
DT 01-MAY-1997 (TREMUREL. 03, CREATED)  
DT 01-MAY-1997 (TREMUREL. 03, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMUREL. 09, LAST SEQUENCE UPDATE)  
DE SDF5 PROTEIN.  
GN SDF5.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97092876.  
RA SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D., NAZAREA M.,  
RA HAMADA T., SATO T., NAKANO T., HONJO T.;  
RT "Characterization of novel secreted and membrane proteins insolated  
RT by the signal sequence trap method";  
RL GENOMICS 37:273-280(1996).  
DR EMBL; D50462; D1009688; -;  
DR MGD; MGI:108078; SDF5.  
SQ SEQUENCE 295 AA; 33501 MW; 329888C8 CRC32;

Query Match 20.7%; Score 6; DB 11; Length 295;  
Best Local Similarity 100.0%; Pred. No. 4.84e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 GOKQGG 266  
|  
|  
|  
|  
|  
Qy 16 GOKQGG 21

RESULT 13  
ID O26544 PRELIMINARY; PRT; 301 AA.  
AC O26544;  
DT 01-NOV-1996 (TREMUREL. 01, CREATED)  
DT 01-NOV-1996 (TREMUREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMUREL. 08, LAST SEQUENCE UPDATE)  
DE G PROTEIN BETA SUBUNIT-LIKE PROTEIN TRANS-SPICED.  
OS SCHISTOSOMA MANSONI (BLOOD FLUKE).  
OC EUKARYOTA; METAZOA; PLATHYHELMINTHES; TREMATODA; DIGenea; STRIGEIDIDA;  
OC SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PUERTO RICO;  
RX MEDLINE; 95394947.  
RA DAVIS R.E., HARDWICK C., TAVERNIER P., HODGSON S., SINGH H.;  
RT "RNA trans-splicing in flatworms. Analysis of trans-spliced mRNAs and  
RT genes in the human parasite, Schistosoma mansoni";  
RL J. BIOL. CHEM. 270:21813-21819(1995).  
DR EMBL; U30261; G1002672; -;  
DR PROSITE; PS00678; G.BETA\_REPEATS; 2.  
DR PFAM; PF00400; G-beta; 5.  
KW REPEAT.  
SQ SEQUENCE 301 AA; 33239 MW; 69273860 CRC32;

Query Match 20.7%; Score 6; DB 5; Length 301;

Best Local Similarity 100.0%; Pred. No. 4.84e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 84 QILLWD 89

|||||

Qy 8 QILLWD 13

#### RESULT 14

ID O58955 PRELIMINARY; PRT; 325 AA.

AC O58955;

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)

DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)

DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)

DE 325AA LONG HYPOTHETICAL PROTEIN.

GN PH1228.

OS PYROCOCUS HORIKOSHII.

OC ARCHAEA; EURYARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCUS.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-OT3;

RX MEDLINE; 98344137.

RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,

RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSOIYAMA A., NAGAI Y.,

RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,

RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,

RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,

RA KIKUCHI H.;

RT "Complete Sequence and Gene Organization of the Genome of a

RT Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii Ot3."

RL DNA RES.5:55-76(1998)

DR EMBL; AP000005; D1031271;

SQ SEQUENCE 325 AA; 36135 MW; C020EDBD CRC32;

Query Match

20.7%; Score 6; DB 1; Length 325;

Best Local Similarity 100.0%; Pred. No. 4.84e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 138 GGFVVD 143

|||||

Qy 20 GGFVVD 25

#### RESULT 15

ID O66891 PRELIMINARY; PRT; 328 AA.

AC O66891;

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)

DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE FLAGELLAR SWITCH PROTEIN FLIG.

GN FLIG.

OS AQUIFEX AEOLICUS.

OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-VF5;

RX MEDLINE; 98196666.

RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,

RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,

RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus."

RL NATURE 392:353-358(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-VF5;

RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,

RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,

RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;

RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; AE000700; G2983250;

KW FLAGELLA.

SQ SEQUENCE 328 AA; 37032 MW; E1C0583A CRC32;

Query Match 20.7%; Score 6; DB 2; Length 328;

Best Local Similarity 100.0%; Pred. No. 4.84e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 256 VVDKNT 261

|||||

Qy 23 VVDKNT 28

Search completed: Sat Aug 28 14:53:02 1999

Job time : 50 secs.

is Page Blank (uspio)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

```
Run on: Sat Aug 28 14:49:42 1999; MasPar time 9.64 Seconds
44.117 Million cell updates/sec
```

Perfect Score: 20

Accounting course:  
Gap 60

Post-processing: Minimum Match 0%

1:part1 2

8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 2.020: Variance 0.732: scale 2.760

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OK

Result No.	Query		Length	DB	ID	Description	Pred. No.
	Score	Match					
1	6	30.0	113	3	R71833	Alternaria alternata	9.42e+01
2	6	30.0	123	5	R26055	Human galactin.	9.42e+01
3	6	30.0	123	5	R37115	Sequence of human pre	9.42e+01
4	6	30.0	153	1	P30951	Herpes Simplex virus-	9.42e+01
5	6	30.0	198	17	R8652	N.meningitidis IM169	9.42e+01
6	6	30.0	198	17	R8652	N.meningitidis IM169	9.42e+01
7	6	30.0	199	17	R8651	N.meningitidis B3032	9.42e+01
8	6	30.0	211	17	R8658	N.meningitidis B283 T	9.42e+01
9	6	30.0	322	4	R23599	Recombinant hematopei	9.42e+01
10	6	30.0	349	4	R23079	Epo-IL-3 Flex, recomb	9.42e+01
11	6	30.0	353	3	R20178	P.glumae PGI lipase s	9.42e+01
12	6	30.0	374	2	R80277	Sequence encoded by 3	9.42e+01
13	6	30.0	426	6	R13036	C. acidovorans stereo	9.42e+01
14	6	30.0	459	27	R36485	Mouse TUB Form I.	9.42e+01
15	6	30.0	505	27	R36486	Mouse TUB Form II.	9.42e+01
16	6	30.0	505	24	U10728	Mouse tub gene produc	9.42e+01

90	5	25.0	115 29	W52883	H. pylori ORF 13ae105	6.89e+02	163	5	25.0	298 33	W58569	Mortierella alpina cy	6.89e+02
91	5	25.0	118 5	R25581	Insulin precursor MI3	6.89e+02	164	5	25.0	315 34	W63616	Hepatitis GB virus E2	6.89e+02
92	5	25.0	119 21	W05714	Alphat1 variable hea	6.89e+02	165	5	25.0	315 3	R13159	2-ketoadonate reduct	6.89e+02
93	5	25.0	120 4	P40076	Sequence of the precu	6.89e+02	166	5	25.0	318 30	W52285	Gallus domesticus cho	6.89e+02
94	5	25.0	120 5	R25582	Insulin precursor MI3	6.89e+02	167	5	25.0	320 28	W27718	S. marcesens Ssp-h2 a	6.89e+02
95	5	25.0	121 5	P20028	Sequence of preprosom	6.89e+02	168	5	25.0	320 28	W27717	S. marcesens Ssp-h1 a	6.89e+02
96	5	25.0	121 18	R97830	Kaposi's sarcoma asso	6.89e+02	169	5	25.0	321 29	W41726	Actinobacillus pleuro	6.89e+02
97	5	25.0	121 17	R93611	HSV-2 strain SB5 asso	6.89e+02	170	5	25.0	321 5	R27118	Fucose dehydrogenase	6.89e+02
98	5	25.0	124 35	W72053	Kaposi's sarcoma asso	6.89e+02	171	5	25.0	321 36	W80603	S. typhimurium PoxR p	6.89e+02
99	5	25.0	124 25	P21889	HSV-2 strain SB5 Cont	6.89e+02	172	5	25.0	329 10	R51284	L-fucose dehydrogenas	6.89e+02
100	5	25.0	124 2	W70018	Partial 4-(4/6) protei	6.89e+02	173	5	25.0	329 26	W29877	Lysophosphatidic acid	6.89e+02
101	5	25.0	124 1	R03587	Interferon-induced vi	6.89e+02	174	5	25.0	331 26	R64769	HSV-2 ISP35 protein.	6.89e+02
102	5	25.0	124 12	R67991	Acetohydroxy acid syn	6.89e+02	175	5	25.0	335 37	W82830	Endothelial MAD inter	6.89e+02
103	5	25.0	127 22	W20547	Insulin precursor MI3	6.89e+02	176	5	25.0	335 11	R57736	Hyperinsensitivity-rela	6.89e+02
104	5	25.0	128 31	W47281	H. pylori secreted or	6.89e+02	177	5	25.0	344 32	W48763	Human B-cell transloc	6.89e+02
105	5	25.0	131 33	W61381	Papillomavirus major	6.89e+02	178	5	25.0	344 24	W18222	Porcine CD34 from clo	6.89e+02
106	5	25.0	131 5	R25579	Human HFGAN72 recepto	6.89e+02	179	5	25.0	345 19	W01098	Human B-cell transloc	6.89e+02
107	5	25.0	133 14	R79376	Insulin precursor MI3	6.89e+02	180	5	25.0	346 32	W60569	Human chromosome 19 d	6.89e+02
108	5	25.0	135 29	W47464	Human ATF-2.	6.89e+02	181	5	25.0	346 37	W82556	Human chromogenin clone	6.89e+02
109	5	25.0	145 32	W60119	Human anhidrotic ecto	6.89e+02	182	5	25.0	348 2	P70508	Corn glycogenin clone	6.89e+02
110	5	25.0	154 39	W95498	Mycobacterium vaccae	6.89e+02	183	5	25.0	349 39	W83928	Human FTHMA-070 parti	6.89e+02
111	5	25.0	159 21	W02177	M. tuberculosis RP-fa	6.89e+02	184	5	25.0	349 1	P81228	Carbonic anhydrase as	6.89e+02
112	5	25.0	165 37	W72887	Residues 1-159 of Bcr	6.89e+02	185	5	25.0	350 21	W02172	Residues 64-413 of Bc	6.89e+02
113	5	25.0	169 32	W42106	Mycobacterium tubercu	6.89e+02	186	5	25.0	352 4	R20794	EHV-4 TK protein.	6.89e+02
114	5	25.0	172 32	W37845	Amino acid sequence o	6.89e+02	187	5	25.0	352 22	W18645	Human tau gene +2 fra	6.89e+02
115	5	25.0	172 37	W80812	Human XAG growth fact	6.89e+02	188	5	25.0	352 25	W17798	Human tau gene +2 fra	6.89e+02
116	5	25.0	172 22	W20609	Amino acid sequence o	6.89e+02	189	5	25.0	359 25	W7798	KN1 alanine scanning	6.89e+02
117	5	25.0	174 19	W04894	H. pylori secreted or	6.89e+02	190	5	25.0	361 10	R51233	Heat resistant alkali	6.89e+02
118	5	25.0	174 19	W04891	Proteinase K resistan	6.89e+02	191	5	25.0	362 39	W95490	M. tuberculosis RP-fa	6.89e+02
119	5	25.0	174 19	W04893	Proteinase K resistan	6.89e+02	192	5	25.0	363 9	R47557	LLTV thymidine kinase	6.89e+02
120	5	25.0	175 19	W04892	Proteinase K resistan	6.89e+02	193	5	25.0	365 25	W18289	Mangifera indica ACC	6.89e+02
121	5	25.0	176 39	W95497	Proteinase K resistan	6.89e+02	194	5	25.0	366 11	R60044	Growth differentiatio	6.89e+02
122	5	25.0	178 31	W53362	M. tuberculosis RP-fa	6.89e+02	195	5	25.0	371 23	W32656	E6AP-binding protein	6.89e+02
123	5	25.0	179 34	W64474	Fragment of the human	6.89e+02	196	5	25.0	374 27	W35370	Hepatitis G virus ant	6.89e+02
124	5	25.0	181 33	W55093	Human secreted protei	6.89e+02	197	5	25.0	376 26	W43043	Herpes simplex virus	6.89e+02
125	5	25.0	186 2	R24190	Streptococcus pneumon	6.89e+02	198	5	25.0	376 26	W43044	Herpes simplex virus	6.89e+02
126	5	25.0	188 27	W24126	Bovine RSV strain A 5	6.89e+02	199	5	25.0	376 26	W43045	Herpes simplex virus	6.89e+02
127	5	25.0	188 16	R87011	Thermococcus protease	6.89e+02	200	5	25.0	376 26	W43045	Xenopus KVLQRL.	6.89e+02
128	5	25.0	193 23	P19743	Peptide sequence.	6.89e+02	201	5	25.0	377 25	W24558	Herpes simplex virus	6.89e+02
129	5	25.0	194 2	P70477	Sugar biosynthesis en	6.89e+02	202	5	25.0	378 21	W12044	Presentin-18 polypept	6.89e+02
130	5	25.0	194 2	P70785	Sequence of human res	6.89e+02	203	5	25.0	380 23	W09406	Mouse Sox-18 polypept	6.89e+02
131	5	25.0	194 29	W47606	Sequence encoding hum	6.89e+02	204	5	25.0	381 13	R59623	Transforming growth f	6.89e+02
132	5	25.0	194 5	R25303	HRV protein 22K.	6.89e+02	205	5	25.0	382 21	W12829	S. avermitilis BCKDH	6.89e+02
133	5	25.0	198 16	R93137	HRV 22K protein.	6.89e+02	206	5	25.0	382 21	W12829	Human recombinant ant	6.89e+02
134	5	25.0	198 1	P90420	Mouse guanylate kinas	6.89e+02	207	5	25.0	384 28	W27303	Human prostate protei	6.89e+02
135	5	25.0	201 38	W81980	Header sequence of mi	6.89e+02	208	5	25.0	386 22	W18664	Fragmented human NF-H	6.89e+02
136	5	25.0	207 7	R33389	Ehrlichia sp. E82.2 p	6.89e+02	209	5	25.0	389 10	R56861	Human placental album m	6.89e+02
137	5	25.0	207 7	R33389	Urease subunit from B	6.89e+02	210	5	25.0	389 6	R33068	Aphanocladium album m	6.89e+02
138	5	25.0	211 38	W85101	Thyroid hormone recep	6.89e+02	211	5	25.0	390 1	R05475	Aphanocladium album m	6.89e+02
139	5	25.0	211 34	W63612	Hepatitis GB virus E2	6.89e+02	212	5	25.0	393 36	W79095	Chloramphenicol resis	6.89e+02
140	5	25.0	211 1	R05220	Antigen GX3276 encode	6.89e+02	213	5	25.0	393 36	W79095	Human secreted protei	6.89e+02
141	5	25.0	212 2	P70055	Fes/fps proto-oncogen	6.89e+02	214	5	25.0	405 31	W56361	Fragment of the human	6.89e+02
142	5	25.0	212 10	R53543	Thyroid hormone recep	6.89e+02	215	5	25.0	405 22	W20805	H. pylori cytoplasmic	6.89e+02
143	5	25.0	220 16	R88269	Thyroid hormone recep	6.89e+02	216	5	25.0	413 21	W02179	Residues 1-413 of Bcr	6.89e+02
144	5	25.0	221 21	W02178	Papilloma virus major	6.89e+02	217	5	25.0	416 14	R81981	Rat prostaglandin I2	6.89e+02
145	5	25.0	228 34	W77740	Residues 1-221 of Bcr	6.89e+02	218	5	25.0	418 11	R60155	Comamonas testosteron	6.89e+02
146	5	25.0	235 11	R57089	Staphylococcus aureus	6.89e+02	219	5	25.0	418 13	R81704	Human NF-Ar transcrip	6.89e+02
147	5	25.0	237 16	R87009	Fasciola hepatica lar	6.89e+02	220	5	25.0	419 27	W30917	DNA gyrase subunit B	6.89e+02
148	5	25.0	242 32	W54943	Hyperthermostable pro	6.89e+02	221	5	25.0	421 13	R74172	RNA polymerase subunit	6.89e+02
149	5	25.0	242 32	W54943	H. pylori cytoplasmic	6.89e+02	222	5	25.0	421 13	R74172	Tyrosinase and tyrosi	6.89e+02
150	5	25.0	251 16	R80192	Rat neurogenin 1 prot	6.89e+02	223	5	25.0	423 6	R33072	Aphanocladium album p	6.89e+02
151	5	25.0	252 35	W72049	Minor ampullate spide	6.89e+02	224	5	25.0	426 21	W02181	pl60 Bcr-Abl.	6.89e+02
152	5	25.0	256 3	P60010	HSV-2 strain SB5 Cont	6.89e+02	225	5	25.0	426 27	W35369	Hepatitis G virus ant	6.89e+02
153	5	25.0	265 17	R93156	Sequence of oncogene	6.89e+02	226	5	25.0	426 31	W59002	Hepatitis G Virus E2	6.89e+02
154	5	25.0	265 28	W46518	Mature inositol-monop	6.89e+02	227	5	25.0	426 13	R74173	Tyrosinase and tyrosi	6.89e+02
155	5	25.0	266 33	W48661	Amino acid sequence o	6.89e+02	228	5	25.0	429 2	R10976	Polyhydroxyalkanoate	6.89e+02
156	5	25.0	267 19	R30522	C. utilis URA3 protei	6.89e+02	229	5	25.0	438 30	W40373	Human breast cancer p	6.89e+02
157	5	25.0	267 34	W77659	Kurthia sp. bioH gene	6.89e+02	230	5	25.0	438 23	W01572	Protein encoded by CH	6.89e+02
158	5	25.0	273 32	W57489	URA3 gene product.	6.89e+02	231	5	25.0	443 34	W01572	Trametes pubescens ph	6.89e+02
159	5	25.0	273 32	W60128	Amino acid sequence o	6.89e+02	232	5	25.0	447 4	R20466	Mouse retinoic acid r	6.89e+02
160	5	25.0	287 5	R30189	M. vaccae antigen GV-	6.89e+02	233	5	25.0	448 13	R66764	Human RAR-beta.	6.89e+02
161	5	25.0	288 35	W60676	UGT1F Exon 1 product.	6.89e+02	234	5	25.0	448 10	R55128	Retinoic acid recepto	6.89e+02
162	5	25.0	290 35	W79398	IMP-18p myo-inositol	6.89e+02	235	5	25.0	448 16	R84725	Murine retinoic acid	6.89e+02
	5	25.0	290 20	W06969	Staphylococcus aureus	6.89e+02				448 17	R55315	Human hepatoma retino	6.89e+02
	5	25.0			M. tuberculosis ORF 3	6.89e+02				448 2	R10548	Murine Retinoic Acid	6.89e+02



236	5	25.0	448	1	P90341	hap (hepatoma) protei	6.89e+02	309	5	25.0	598	32	W59461	Microbial L-alpha-gly	6.89e+02
237	5	25.0	449	30	W52286	Gallus domesticus cho	6.89e+02	310	5	25.0	598	32	W59462	Microbial L-alpha-gly	6.89e+02
238	5	25.0	452	5	R27797	EPSP synanth wild ty	6.89e+02	311	5	25.0	598	32	W59463	Microbial L-alpha-gly	6.89e+02
239	5	25.0	454	2	R10182	Recombinant human gam	6.89e+02	312	5	25.0	599	21	W03692	Corn acetohydroxy aci	6.89e+02
240	5	25.0	454	16	R84723	Human Retinoic acid R	6.89e+02	313	5	25.0	599	32	W59496	Microbial L-alpha-gly	6.89e+02
241	5	25.0	454	2	R10550	Human Retinoic acid R	6.89e+02	314	5	25.0	599	21	W16598	Corn acetohydroxy aci	6.89e+02
242	5	25.0	457	27	W25789	Acromonium cellulolyt	6.89e+02	315	5	25.0	599	21	W16599	Corn acetohydroxy aci	6.89e+02
243	5	25.0	458	16	R84726	Murine retinoic acid	6.89e+02	316	5	25.0	599	21	W16597	Corn acetohydroxy aci	6.89e+02
244	5	25.0	458	2	R10549	Skin-specific murine	6.89e+02	317	5	25.0	603	3	R12228	Human TR2-11 DNA bind	6.89e+02
245	5	25.0	460	17	R20465	Mouse retinoic acid r	6.89e+02	318	5	25.0	604	3	W80480	Islet cell antibody a	6.89e+02
246	5	25.0	460	17	R88651	N.meningitidis IM2169	6.89e+02	319	5	25.0	604	2	R06463	Derived protein of cl	6.89e+02
247	5	25.0	462	39	W73623	Human secreted protei	6.89e+02	320	5	25.0	606	1	P90110	Polypeptide having L-	6.89e+02
248	5	25.0	466	31	W59165	L. lactis WGL316 gadB	6.89e+02	321	5	25.0	607	32	W59465	Microbial L-alpha-gly	6.89e+02
249	5	25.0	467	3	R12227	Human TR2-9 DNA bindi	6.89e+02	322	5	25.0	607	32	W59456	Microbial L-alpha-gly	6.89e+02
250	5	25.0	476	19	W02616	Wheat adenylosuccinat	6.89e+02	323	5	25.0	608	26	W14645	N. meningitidis IM216	6.89e+02
251	5	25.0	478	13	R74174	Chloroplast transit p	6.89e+02	324	5	25.0	608	17	R88650	N.meningitidis IM2169	6.89e+02
252	5	25.0	478	9	R47586	Human bone formatio-	6.89e+02	325	5	25.0	609	6	R32131	CMV Colburn region po	6.89e+02
253	5	25.0	478	20	W06539	Human bone morphogeni	6.89e+02	326	5	25.0	610	38	W65769	Soybean phragmoplasti	6.89e+02
254	5	25.0	482	32	W60234	Pelargonium l-aminocy	6.89e+02	327	5	25.0	610	38	W65770	Soybean phragmoplasti	6.89e+02
255	5	25.0	482	32	W60233	Pelargonium l-aminocy	6.89e+02	328	5	25.0	620	36	W72129	HSV-2 strain SB5 Cont	6.89e+02
256	5	25.0	482	22	W09878	ACC synthase GAC-1.	6.89e+02	329	5	25.0	620	36	W72129	HSV-2 strain SB5 Cont	6.89e+02
257	5	25.0	483	1	P90598	TR2-5 DNA clone.	6.89e+02	330	5	25.0	623	38	W88707	Secreted protein enco	6.89e+02
258	5	25.0	483	3	R12226	Human TR2-5 androgen	6.89e+02	331	5	25.0	633	3	R20030	Bacillus thuringiens	6.89e+02
259	5	25.0	485	2	R11919	Reverse transcriptase	6.89e+02	332	5	25.0	633	35	W75775	Amino acid sequence o	6.89e+02
260	5	25.0	486	2	R08034	Unique sequence fragm	6.89e+02	333	5	25.0	633	35	W75774	Amino acid sequence o	6.89e+02
261	5	25.0	492	24	W27613	Brevibacterium flavum	6.89e+02	334	5	25.0	636	12	R64767	HSV-2 protease, ICP35	6.89e+02
262	5	25.0	493	3	R15504	Zucchini ACC synthase	6.89e+02	335	5	25.0	637	33	W64388	HSV-2 protease, ICP35	6.89e+02
263	5	25.0	493	29	W47311	Zucchini ACC synthase	6.89e+02	336	5	25.0	638	31	W54053	A. thaliana SGR prote	6.89e+02
264	5	25.0	493	29	W47310	Protein encoded by zu	6.89e+02	337	5	25.0	638	31	W54055	AHAS clone.	6.89e+02
265	5	25.0	493	3	R15505	Zucchini ACC synthase	6.89e+02	338	5	25.0	638	1	R05588	AHAS clone.	6.89e+02
266	5	25.0	493	5	R25406	ACC synthetase.	6.89e+02	339	5	25.0	638	6	R32006	Acetohydroxy acid syn	6.89e+02
267	5	25.0	494	29	W47312	Zucchini ACC synthase	6.89e+02	340	5	25.0	638	6	R32006	Herbicide resistant A	6.89e+02
268	5	25.0	494	3	R15963	Zucchini ACC synthase	6.89e+02	341	5	25.0	638	6	R32007	Acetohydroxy acid syn	6.89e+02
269	5	25.0	501	4	R23577	Phospholipase D-K fro	6.89e+02	342	5	25.0	638	13	R71017	Herbicide sensitive A	6.89e+02
270	5	25.0	508	3	R13435	Terminal deoxynucleot	6.89e+02	343	5	25.0	638	13	W54054	HSV-2 protease.	6.89e+02
271	5	25.0	513	2	R07311	Human Bone Morphogene	6.89e+02	344	5	25.0	638	6	R32008	AHAS clone.	6.89e+02
272	5	25.0	513	7	R36735	Human BMP-6.	6.89e+02	345	5	25.0	642	36	W57979	Herbicide sensitive A	6.89e+02
273	5	25.0	513	6	R32904	Human BMP-6.	6.89e+02	346	5	25.0	642	36	W72124	Herbicide sensitive A	6.89e+02
274	5	25.0	513	9	R47264	Pre-pro BMP6.	6.89e+02	347	5	25.0	643	39	W89300	Amino acid sequence o	6.89e+02
275	5	25.0	517	1	R06260	Human acetylcholine r	6.89e+02	348	5	25.0	643	19	W03738	HSV-2 strain SB5 Cont	6.89e+02
276	5	25.0	522	27	W24122	Pyrococcus furiosus p	6.89e+02	349	5	25.0	653	1	R06723	Human rchd502 protein	6.89e+02
277	5	25.0	527	36	W75776	Mycobacterium tubercu	6.89e+02	350	5	25.0	653	1	R06723	rchd502 gene product.	6.89e+02
278	5	25.0	528	38	W81457	Mycobacterium tubercu	6.89e+02	351	5	25.0	654	27	W24129	Acromobacter proteas	6.89e+02
279	5	25.0	532	35	W49700	Canine ICAM-R polypep	6.89e+02	352	5	25.0	657	36	W72192	Pyrococcus furiosus p	6.89e+02
280	5	25.0	533	37	W83317	Human flavin-contains	6.89e+02	353	5	25.0	659	8	R39308	HSV-2 strain SB5 Cont	6.89e+02
281	5	25.0	535	35	W49699	Mouse Lrp5 protein fr	6.89e+02	354	5	25.0	659	25	W35827	Mexico strain HEV ORF	6.89e+02
282	5	25.0	536	30	W53105	Human flavin-contains	6.89e+02	355	5	25.0	659	25	W35827	Protease.	6.89e+02
283	5	25.0	537	5	R25534	T. gondii chorismate	6.89e+02	356	5	25.0	659	27	W24121	Hepatitis E virus Mex	6.89e+02
284	5	25.0	540	24	W12529	Toxoplasma gondii 54k	6.89e+02	357	5	25.0	659	8	R38787	Thermococcus proteas	6.89e+02
285	5	25.0	541	22	W05767	Human brain-specific	6.89e+02	358	5	25.0	660	37	W81520	HEV ORF2 protein.	6.89e+02
286	5	25.0	559	37	W68562	Presenilin homologue.	6.89e+02	359	5	25.0	660	19	R91814	Hepatitis E virus (HE	6.89e+02
287	5	25.0	559	37	W68562	Pseudomonas species p	6.89e+02	360	5	25.0	660	17	R96090	Hepatitis E virus str	6.89e+02
288	5	25.0	559	2	P70580	Plasmid pTUB616 heat	6.89e+02	361	5	25.0	660	10	R51265	Hepatitis E virus (Me	6.89e+02
289	5	25.0	559	23	W19630	Streptomyces venezuel	6.89e+02	362	5	25.0	663	36	W49872	HEV strain protein en	6.89e+02
290	5	25.0	559	2	R10882	Polyhydroxyalkanoate	6.89e+02	363	5	25.0	672	13	R71325	Thermotoga maritima M	6.89e+02
291	5	25.0	560	12	R71294	Human glycoprotein v.	6.89e+02	364	5	25.0	674	26	W36522	Poly-beta-hydroxyalka	6.89e+02
292	5	25.0	572	2	P70579	Plasmid pTUB613 heat	6.89e+02	365	5	25.0	674	33	W63489	Human adducin-like 70	6.89e+02
293	5	25.0	578	31	W53551	SIRE-1 genomic clone	6.89e+02	366	5	25.0	677	38	W85608	A. thaliana protein c	6.89e+02
294	5	25.0	581	26	W33355	Human KVLQTL associat	6.89e+02	367	5	25.0	677	38	W85608	Secreted protein clon	6.89e+02
295	5	25.0	581	25	W30038	Human KVLQTL associat	6.89e+02	368	5	25.0	680	26	W49868	Thermotoga maritima b	6.89e+02
296	5	25.0	582	11	R42306	Phospholipase D-P	6.89e+02	369	5	25.0	680	26	W49868	Thermotoga maritima 6	6.89e+02
297	5	25.0	587	17	R88656	N.meningitidis B283 T	6.89e+02	370	5	25.0	680	26	W49868	Thermotoga maritima e	6.89e+02
298	5	25.0	589	6	R32190	Sequence encoded by t	6.89e+02	371	5	25.0	691	4	R24024	Thermotoga maritima	6.89e+02
299	5	25.0	589	2	R10681	Polyhydroxybutyrate p	6.89e+02	372	5	25.0	695	31	W53895	Sequence encoded by t	6.89e+02
300	5	25.0	590	1	P96204	Human muscarinic acet	6.89e+02	373	5	25.0	700	2	R13354	SIRE-1 genomic clone	6.89e+02
301	5	25.0	591	37	W72912	Mycobacterium tubercu	6.89e+02	374	5	25.0	706	33	W69846	Cellular Receptor 2 l	6.89e+02
302	5	25.0	595	30	W53347	Nephila clavipes spid	6.89e+02	375	5	25.0	710	33	W69846	Amino acid sequence o	6.89e+02
303	5	25.0	595	3	R14309	N.clavipes dragline s	6.89e+02	376	5	25.0	724	13	R71383	Amphiphilic semaphorin	6.89e+02
304	5	25.0	596	15	R90715	Thermotoga maritima	6.89e+02	377	5	25.0	765	17	R38150	Human NF-AT transcrip	6.89e+02
305	5	25.0	598	32	W59460	Microbial L-alpha-gly	6.89e+02	378	5	25.0	767	17	R73016	Acetobacter diazotifia	6.89e+02
306	5	25.0	598	32	W59458	Microbial L-alpha-gly	6.89e+02	379	5	25.0	778	21	W15489	Fish protamine gene t	6.89e+02
307	5	25.0	598	32	W59464	Microbial L-alpha-gly	6.89e+02	380	5	25.0	784	27	W34179	Pemphigus foliaceus a	6.89e+02
308	5	25.0	598	32	W59459	Microbial L-alpha-gly	6.89e+02	381	5	25.0	789	3	R15143	Human GC binding prot	6.89e+02
											790	26	W31528	Haemonchus contortus	6.89e+02
											796	20	W08087	Human anti-apoptotic	6.89e+02
														Infective bursa prote	6.89e+02

382	5	25.0	797 31	W56368	Amino acid sequence o	6.89e+02						SPV4 non-structural p	6.89e+02
383	5	25.0	797 31	W56377	Human GLYT-2 transpor	6.89e+02	455	25.0	2431 5	R25138		Photorehabdus luminesc	6.89e+02
384	5	25.0	797 31	W56384	Human GLYT-2 transpor	6.89e+02	456	25.0	2504 25	W17871		Toxin Tcba, encoded b	6.89e+02
385	5	25.0	797 31	W56375	Human GLYT-2 transpor	6.89e+02	457	25.0	2504 31	W56557		Rat IP3 receptor..	6.89e+02
386	5	25.0	797 31	W56380	Human GLYT-2 transpor	6.89e+02	458	25.0	2670 15	R88125		Human IP3 receptor..	6.89e+02
387	5	25.0	797 31	W56378	Human GLYT-2 transpor	6.89e+02	459	25.0	2670 15	R88126		Human type 1 inositol	6.89e+02
388	5	25.0	797 31	W56379	Human GLYT-2 transpor	6.89e+02	460	25.0	2695 20	W07632		Human type 3 inositol	6.89e+02
389	5	25.0	797 31	W56374	Human GLYT-2 transpor	6.89e+02	461	25.0	2701 18	R99850		Human T cell inositol	6.89e+02
390	5	25.0	797 31	W56374	Human GLYT-2 transpor	6.89e+02	462	25.0	2713 19	W00168		Fragment HGI1775 of a	6.89e+02
391	5	25.0	797 31	W56386	Human GLYT-2 transpor	6.89e+02	463	25.0	2723 31	W56448		Fragment HGI1775 of a	6.89e+02
392	5	25.0	797 31	W56373	Human GLYT-2 transpor	6.89e+02	464	25.0	2749 3	R13887		Inositol-3-phosphate	6.89e+02
393	5	25.0	797 31	W56376	Human GLYT-2 transpor	6.89e+02	465	25.0	2861 20	W06537		Hepatitis GB virus is	6.89e+02
394	5	25.0	797 31	W56383	Human GLYT-2 transpor	6.89e+02	466	25.0	2860 34	W03611		Hepatitis GB virus E2	6.89e+02
395	5	25.0	797 31	W56369	Human GLYT-2 transpor	6.89e+02	467	25.0	2873 39	W84542		Hepatitis G virus PNF	6.89e+02
396	5	25.0	797 31	W56372	Amino acid sequence o	6.89e+02	468	25.0	2873 28	R78559		Hepatitis virus clone	6.89e+02
397	5	25.0	797 31	W56385	Human GLYT-2 transpor	6.89e+02	469	25.0	2873 28	W34983		Hepatitis G virus PNF	6.89e+02
398	5	25.0	797 31	W56382	Human GLYT-2 transpor	6.89e+02	470	25.0	2873 18	R30796		Hepatitis G virus PNF	6.89e+02
399	5	25.0	802 35	W59757	Acetobacter xylinum b	6.89e+02	471	25.0	2873 34	W60054		HGV isolate PNF 2161	6.89e+02
400	5	25.0	808 5	R25589	RING4 antigenic pepti	6.89e+02	472	25.0	2873 36	W80148		HGV-PNF2161 variant s	6.89e+02
401	5	25.0	818 35	W72035	HSV-2 strain SB5 Cont	6.89e+02	473	25.0	2873 31	W56441		Fragment HGI606 of a	6.89e+02
402	5	25.0	831 12	R64273	T. flavus DNA-polymer	6.89e+02	474	25.0	2905 20	W06536		Hepatitis GB virus is	6.89e+02
403	5	25.0	831 37	W80428	DNA polymerase enzyme	6.89e+02	475	25.0	2910 39	W89458		Hepatitis G virus var	6.89e+02
404	5	25.0	831 37	W79961	Thermus flavus DNA po	6.89e+02	476	25.0	2910 36	R80182		Hepatitis G virus (HG	6.89e+02
405	5	25.0	848 34	W29677	Homo sapiens clone CS	6.89e+02	477	25.0	2910 18	R90797		HGV-JC variant polypr	6.89e+02
406	5	25.0	850 31	W57445	A. thermophilum therm	6.89e+02	478	25.0	2910 34	W76088		HGV variant JC protei	6.89e+02
407	5	25.0	856 37	W75221	Human secreted protei	6.89e+02	479	25.0	2963 31	R75766		Hepatitis G virus clo	6.89e+02
408	5	25.0	881 36	W72174	HSV-2 strain SB5 Cont	6.89e+02	480	25.0	2963 31	W56444		Fragment HGI1789 of a	6.89e+02
409	5	25.0	881 35	W72002	HSV-2 strain SB5 Cont	6.89e+02	481	25.0	2969 31	W56446		Fragment HGI2141 of a	6.89e+02
410	5	25.0	887 25	W19856	Rat semaphorin 2.	6.89e+02	482	25.0	2969 31	W56449		Fragment K3732(2Y) of	6.89e+02
411	5	25.0	888 25	W19857	Human semaphorin 2.	6.89e+02	483	25.0	2969 31	W56443		Fragment HGI1741 of a	6.89e+02
412	5	25.0	894 39	W68610	WprA protein.	6.89e+02	484	25.0	2969 31	W56442		Fragment HGI1668 of a	6.89e+02
413	5	25.0	904 37	W63511	Human DNAX toll-like	6.89e+02	485	25.0	2970 31	W56445		Fragment HGI1916 of a	6.89e+02
414	5	25.0	906 27	W36795	Novel human protein.	6.89e+02	486	25.0	2985 32	W56643		DNA-dependent protein	6.89e+02
415	5	25.0	912 18	R78888	Bacterial transferrin	6.89e+02	487	25.0	3011 16	R95020		Hepatitis GB virus (H	6.89e+02
416	5	25.0	914 25	W46800	Spinocerebellar ataxi	6.89e+02	488	25.0	3119 36	W72204		HSV-2 strain SB5 Cont	6.89e+02
417	5	25.0	921 20	W06077	Drosophila TATA-bind	6.89e+02	489	25.0	3224 31	W52435		Human Nup358 protein.	6.89e+02
418	5	25.0	921 23	W55028	TATA-binding protein	6.89e+02	490	25.0	3567 8	R44431		eryA region polypepti	6.89e+02
419	5	25.0	921 11	R56407	TATA-binding protein-	6.89e+02	491	25.0	3722 2	R10145		Cephalosporin antibio	6.89e+02
420	5	25.0	969 8	R41662	Paired basic amino ac	6.89e+02	492	25.0	3910 12	R66462		ALI-1 (acute lymphocy	6.89e+02
421	5	25.0	973 21	W11853	Hepadnavirus receptor	6.89e+02	493	25.0	3910 7	R38470		ALL-1 protein.	6.89e+02
422	5	25.0	997 10	R43630	Mannuronan C-5-epimer	6.89e+02	494	25.0	3960 32	W56642		Protein kinase cataly	6.89e+02
423	5	25.0	1025 18	W03185	Rice Xa21 disease res	6.89e+02	495	25.0	3969 10	R52971		Product of the CDNA e	6.89e+02
424	5	25.0	1048 27	W72177	Human cytomegalovirus	6.89e+02	496	25.0	4472 19	R92741		Virulence gene cluste	6.89e+02
425	5	25.0	1048 4	R20637	Human Cytomegalovirus	6.89e+02	497	25.0	4672 30	W52845		A. mediterranei rifam	6.89e+02
426	5	25.0	1085 2	R16104	P450 17-alpha/P450 BM	6.89e+02	498	25.0	4630 23	W19629		Streptomyces venezuel	6.89e+02
427	5	25.0	1085 3	R61048	A.nidulans phosphenol	6.89e+02	499	25.0	4655 17	R97209		Human placental calci	6.89e+02
428	5	25.0	1088 16	R86636	Plasmid pASK75 open r	6.89e+02	500	25.0	4655 29	W43313		Human kidney calcium	6.89e+02
429	5	25.0	1122 12	R49297	Cytadhesin protein.	6.89e+02	501	25.0	4655 29	W43313		Human kidney calcium	6.89e+02
430	5	25.0	1233 29	W41935	Nudauralia beta-like	6.89e+02	502	25.0	4655 29	W43312		Human placental calci	6.89e+02
431	5	25.0	1233 29	W34536	Nudauralia beta virus	6.89e+02	503	25.0	4655 17	R97210		Human kidney calcium	6.89e+02
432	5	25.0	1277 10	R52701	Plasmid pASK60-Strep	6.89e+02	504	25.0	4655 17	R97211		Human parathyroid cal	6.89e+02
433	5	25.0	1284 1	P81187	Sequence encoded by a	6.89e+02	505	25.0	4655 29	W43314		Human parathyroid cal	6.89e+02
434	5	25.0	1297 30	W40200	Infected cell protein	6.89e+02	506	25.0	4655 29	W43311		Human calcium sensor	6.89e+02
435	5	25.0	1298 37	W08010	The amino acid sequen	6.89e+02	507	25.0	10 38	W67452		Human gamma-sarcoglyc	4.33e+03
436	5	25.0	1312 30	W33807	Human ataxin-2.	6.89e+02	508	25.0	10 3	P61448		Luteinising hormone r	4.33e+03
437	5	25.0	1313 33	W60213	Spinocerebellar ataxi	6.89e+02	510	25.0	12 39	P61447		Peptide 63-74 of hspL	4.33e+03
438	5	25.0	1528 18	R95233	Manduca sexta Bacilli	6.89e+02	511	25.0	12 39	W73567		Graminae pollen aller	4.33e+03
439	5	25.0	1545 27	W23362	Human canalicular mul	6.89e+02	512	25.0	12 36	W76451		Peptide from humanise	4.33e+03
440	5	25.0	1545 31	W55966	Human canalicular mul	6.89e+02	513	25.0	13 39	W86097		Peptide from humanise	4.33e+03
441	5	25.0	1591 38	W83308	Mature LRP5 protein.	6.89e+02	514	25.0	13 39	W86098		Peptide from humanise	4.33e+03
442	5	25.0	1614 38	W83312	Mouse Lrp5 protein.	6.89e+02	515	25.0	13 39	W86079		Peptide from humanise	4.33e+03
443	5	25.0	1615 38	W83309	LRP5 protein from the	6.89e+02	516	25.0	13 39	W86080		Peptide from humanise	4.33e+03
444	5	25.0	1721 30	W52847	A. mediterranei rifam	6.89e+02	517	25.0	15 39	W73843		M. tuberculosis antig	4.33e+03
445	5	25.0	1738 10	R50971	Norwalk virus strain	6.89e+02	518	25.0	15 39	W87801		N-terminal sequence o	4.33e+03
446	5	25.0	1841 26	W22605	Tyactone synthase OR	6.89e+02	519	25.0	15 39	W73754		M. tuberculosis antig	4.33e+03
447	5	25.0	1844 25	W18302	Photorehabdus luminesc	6.89e+02	520	25.0	15 39	W73732		M. tuberculosis antig	4.33e+03
448	5	25.0	1844 31	W56558	Toxin TcbaII, encoded	6.89e+02	521	25.0	15 39	W73733		M. tuberculosis antig	4.33e+03
449	5	25.0	1896 36	W72095	HSV-2 strain SB5 Cont	6.89e+02	522	25.0	15 39	W73844		M. tuberculosis der	4.33e+03
450	5	25.0	1996 26	W23717	Platenolide synthase	6.89e+02	523	25.0	15 39	W95431		Antiviral peptide der	4.33e+03
451	5	25.0	1996 26	W22607	Platenolide synthase	6.89e+02	524	25.0	15 39	W73842		M. tuberculosis antig	4.33e+03
452	5	25.0	2289 25	W14987	Protein derived from	6.89e+02	525	25.0	16 39	W73367		Human PAR-AH protein	4.33e+03
453	5	25.0	2414 30	W40057	Cellular transcriptio	6.89e+02	526	25.0	17 38	W81649		Mouse liyor-1 145p	4.33e+03
454	5	25.0	2414 14	R84882	Transcription factor	6.89e+02	527	25.0	18 21	W09763		Endothelin antagonist	4.33e+03

528	4	20.0	20 38	W67515	IGF-1/IGFBP inhibitor	4.33e+03	601	121 39	W89484	Hepatitis G virus ant	4.33e+03
529	4	20.0	24 38	W83434	PR-R major amino acid	4.33e+03	602	122 6	R30772	huxCD3v9, humanised m	4.33e+03
530	4	20.0	24 38	W86028	Peptide epitope 4 of	4.33e+03	603	123 39	R30773	Consensus humanised m	4.33e+03
531	4	20.0	24 39	W73563	Peptide 51-74 of hspL	4.33e+03	604	123 39	W87627	Dirofilaria immitis a	4.33e+03
532	4	20.0	24 39	W80960	PS116 antigen.	4.33e+03	605	123 38	W86956	PS116 antigen.	4.33e+03
533	4	20.0	24 39	W73539	PR-R major fragment.	4.33e+03	606	123 38	W86808	Variable Heavy domain	4.33e+03
534	4	20.0	25 36	W75011	Human secreted protei	4.33e+03	607	123 38	W86024	UT116 polypeptide con	4.33e+03
535	4	20.0	26 39	W73340	PR-R minor fragment.	4.33e+03	608	123 38	W86807	Variable Heavy domain	4.33e+03
536	4	20.0	26 38	W52771	Human growth hormone	4.33e+03	609	124 39	W73562	Human group II secret	4.33e+03
537	4	20.0	26 38	W83435	Human growth hormone	4.33e+03	610	126 39	W95183	Mouse MIN6-derived G-	4.33e+03
538	4	20.0	26 38	W85677	Secreted protein enco	4.33e+03	611	127 11	R62085	Tapeworm antigenic pr	4.33e+03
539	4	20.0	26 38	W85676	Human growth factor r	4.33e+03	612	130 29	W40121	Murine monoclonal ant	4.33e+03
540	4	20.0	26 21	W16426	Human growth hormone	4.33e+03	613	131 39	W81638	Mouse liyor-1 protein	4.33e+03
541	4	20.0	27 38	W52760	Human growth hormone	4.33e+03	614	131 38	W86277	Blasx output of huma	4.33e+03
542	4	20.0	27 21	W16415	Human growth factor r	4.33e+03	615	135 18	R90685	Humanized 5C7.29 anti	4.33e+03
543	4	20.0	28 38	W52786	Human growth hormone	4.33e+03	616	135 38	W82266	T. pallidum 17 kDa li	4.33e+03
544	4	20.0	30 3	R15563	Immunopeptide #3 deri	4.33e+03	617	140 6	R29578	CLN-IgG gamma.	4.33e+03
545	4	20.0	30 38	W89959	Antigen 3 from cluste	4.33e+03	618	141 39	W73652	M. tuberculosis antig	4.33e+03
546	4	20.0	31 38	W73491	Cellobiose phosphoryl	4.33e+03	619	141 39	W73762	M. tuberculosis antig	4.33e+03
547	4	20.0	44 39	W82293	Duodenal lumen to bra	4.33e+03	620	142 39	W67919	Human secreted protei	4.33e+03
548	4	20.0	58 18	R89440	CD7 transmembrane dom	4.33e+03	621	143 39	W67918	Human secreted protei	4.33e+03
549	4	20.0	58 31	W28043	Staphylococcus aureus	4.33e+03	622	145 39	W87593	Antibody 11D10 light	4.33e+03
550	4	20.0	58 9	R48131	BPTI (131, 15R, 16A,	4.33e+03	623	149 39	W82406	Human NMLY-6 analogue	4.33e+03
551	4	20.0	58 9	R48129	BPTI (11E, 13I, 15R,	4.33e+03	624	153 17	R87527	Mel-linked mlga gene	4.33e+03
552	4	20.0	64 16	R78151	Zebrafish sonic hedge	4.33e+03	625	153 10	R54634	Consensus sequence of	4.33e+03
553	4	20.0	67 38	W85093	Thyroid hormone recep	4.33e+03	626	153 38	W73358	S. colwelliana Mlga p	4.33e+03
554	4	20.0	73 39	W89478	Hepatitis G virus ant	4.33e+03	627	160 39	W73637	S. pneumoniae lac pr	4.33e+03
555	4	20.0	78 39	W95484	Human-derived RT3 pha	4.33e+03	628	161 39	W89233	Tumour necrosis inhib	4.33e+03
556	4	20.0	91 39	W89239	Protein SEQ ID NO:86	4.33e+03	629	163 38	W86217	Human VRF-2 truncated	4.33e+03
557	4	20.0	91 31	W58899	Plasmid PCSJ bacterio	4.33e+03	630	163 39	W89805	Staphylococcus aureus	4.33e+03
558	4	20.0	92 39	W95480	Mouse derived RT3 pha	4.33e+03	631	168 37	W69496	G. oxydans Cytochrome	4.33e+03
559	4	20.0	92 39	W95476	Mouse derived RT3 pha	4.33e+03	632	168 38	W86216	Human VRF-2 truncated	4.33e+03
560	4	20.0	93 39	W89462	Hepatitis G virus epi	4.33e+03	633	170 39	W81555	Bacillus subtilis rib	4.33e+03
561	4	20.0	94 39	W95316	Chlamydial major oute	4.33e+03	634	170 1	R02226	Myelin basic protein	4.33e+03
562	4	20.0	94 39	W89461	Hepatitis G virus epi	4.33e+03	635	173 38	R86215	Human VRF-2 truncated	4.33e+03
563	4	20.0	94 39	W89240	Protein SEQ ID NO:87	4.33e+03	636	173 39	W89767	Staphylococcus aureus	4.33e+03
564	4	20.0	95 38	W85908	PrP 81 light chain va	4.33e+03	637	174 38	W86281	Cancer marker MUC-B1	4.33e+03
565	4	20.0	95 25	W35098	Product of target DNA	4.33e+03	638	175 39	W84333	GP4 protein of PRRSV	4.33e+03
566	4	20.0	96 39	W67917	Human secreted protei	4.33e+03	639	178 39	W84325	GP4 protein of PRRSV	4.33e+03
567	4	20.0	97 39	W73869	M. tuberculosis antig	4.33e+03	640	178 38	W86214	Human VRF-2 truncated	4.33e+03
568	4	20.0	97 39	W73759	M. tuberculosis antig	4.33e+03	641	178 39	W84324	GP4 protein of PRRSV	4.33e+03
569	4	20.0	97 39	W73653	M. tuberculosis antig	4.33e+03	642	178 39	W84326	GP4 protein of PRRSV	4.33e+03
570	4	20.0	97 39	W86160	P. americanus antilre	4.33e+03	643	178 39	W84327	GP4 protein of PRRSV	4.33e+03
571	4	20.0	97 39	W73763	M. tuberculosis antig	4.33e+03	644	178 39	W84328	GP4 protein of PRRSV	4.33e+03
572	4	20.0	99 39	W73660	M. tuberculosis antig	4.33e+03	645	178 39	W84329	GP4 protein of PRRSV	4.33e+03
573	4	20.0	99 39	W73663	M. tuberculosis antig	4.33e+03	646	181 38	W89810	Protein encoded by cl	4.33e+03
574	4	20.0	99 39	W73766	M. tuberculosis antig	4.33e+03	647	185 38	W86235	Human VRF (VEGF-relat	4.33e+03
575	4	20.0	99 39	W73770	M. tuberculosis antig	4.33e+03	648	188 39	W89277	Rat mast cell functio	4.33e+03
576	4	20.0	99 39	W73773	M. tuberculosis antig	4.33e+03	649	188 38	W89278	Protein encoded by cl	4.33e+03
577	4	20.0	99 39	W73856	M. tuberculosis antig	4.33e+03	650	189 38	W73224	Bovine interferon-alp	4.33e+03
578	4	20.0	101 39	W89238	Protein SEQ ID NO:85	4.33e+03	651	189 38	W83895	Bovine interferon-alp	4.33e+03
579	4	20.0	102 39	W95486	Human-derived RT3 pha	4.33e+03	652	190 39	W89897	Antigen 3 from cluste	4.33e+03
580	4	20.0	103 38	W88536	Secreted protein enco	4.33e+03	653	191 38	W85579	Human N-methyl-D-aspa	4.33e+03
581	4	20.0	106 39	W89235	Protein SEQ ID NO:82	4.33e+03	654	193 38	W73365	Macaque PAF-AH protei	4.33e+03
582	4	20.0	106 38	W88840	Polypeptide fragment	4.33e+03	655	194 38	W85110	Thyroid hormone recep	4.33e+03
583	4	20.0	106 22	W20984	H. pylori cytoplasmic	4.33e+03	656	195 39	W88500	Human stomach carcino	4.33e+03
584	4	20.0	109 39	W89236	Protein SEQ ID NO:83	4.33e+03	657	202 38	W88629	Secreted protein enco	4.33e+03
585	4	20.0	109 38	W85909	PrP 28 light chain va	4.33e+03	658	202 38	W89831	Protein encoded by cl	4.33e+03
586	4	20.0	109 38	W85910	PrP 37 light chain va	4.33e+03	659	203 39	W89759	Staphylococcus aureus	4.33e+03
587	4	20.0	109 39	W89237	Protein SEQ ID NO:84	4.33e+03	660	206 39	W95182	Mouse MIN6-derived G-	4.33e+03
588	4	20.0	109 25	W18271	PrP 37 light chain va	4.33e+03	661	206 38	W86202	Human VEGF-related fa	4.33e+03
589	4	20.0	110 39	W95224	PT toxin beta-subunit	4.33e+03	662	210 39	W86307	Kidney injury associa	4.33e+03
590	4	20.0	114 39	W87708	A cysteine rich solub	4.33e+03	663	210 38	W89990	Expressed antigen for	4.33e+03
591	4	20.0	117 18	R96800	Human histiocyte-secr	4.33e+03	664	210 3	R13499	P.denitrificans COB H	4.33e+03
592	4	20.0	117 39	W86135	Protein sequence of h	4.33e+03	665	211 39	W89225	Tumour necrosis facto	4.33e+03
593	4	20.0	117 39	W86139	Protein sequence of m	4.33e+03	666	212 24	W15932	Antibody 7G12 heavy c	4.33e+03
594	4	20.0	117 39	W86137	Protein sequence of d	4.33e+03	667	214 39	W82684	L. corniculatum type	4.33e+03
595	4	20.0	119 38	W73503	Antibody 15D3 heavy c	4.33e+03	668	215 39	W88238	Hiv-1 co-receptor CCR	4.33e+03
596	4	20.0	121 39	W86118	Murine 340 Vh amino a	4.33e+03	669	218 38	W89832	Protein encoded by cl	4.33e+03
597	4	20.0	121 39	W86120	Protein sequence of h	4.33e+03	670	222 30	W52131	A novel human mRNA ed	4.33e+03
598	4	20.0	121 39	W86125	Protein sequence of m	4.33e+03	671	222 24	W15934	Antibody 3G2 heavy ch	4.33e+03
599	4	20.0	121 39	W86122	Protein sequence of d	4.33e+03	672	223 6	R32841	VH NQ10/12.5-VK NQ10/	4.33e+03
600	4	20.0	121 39	W86124	Protein sequence of a	4.33e+03	673	225 2	R05710	TRY40.	4.33e+03

674	4	20.0	229 38	W89855	Antigen 1 from cluste	4.33e+03	4	20.0	339 6	R33556	Sequence of Lol p Ib.	4.33e+03
675	4	20.0	230 11	R59801	Fusarium solani pisi	4.33e+03	4	20.0	341 2	R06522	Tapetum-specific prom	4.33e+03
676	4	20.0	230 38	W89947	Antigen 2 from cluste	4.33e+03	4	20.0	342 38	W89085	Polypeptide fragment	4.33e+03
677	4	20.0	230 39	W89978	Protein encoded by cl	4.33e+03	4	20.0	343 39	W87713	An active acyltransfe	4.33e+03
678	4	20.0	230 24	W27088	Mouse monoclonal anti	4.33e+03	4	20.0	344 38	W83139	BT20 protein.	4.33e+03
679	4	20.0	232 38	W67489	S. cerevisiae T1H3 po	4.33e+03	4	20.0	344 7	R34579	Phase I5 integrase pr	4.33e+03
680	4	20.0	234 11	R60794	Kawasaki syndrome tox	4.33e+03	4	20.0	347 38	R90057	Rat GlcAT-P protein.	4.33e+03
681	4	20.0	236 39	W89788	Staphylococcus aureus	4.33e+03	4	20.0	349 21	W12719	PhID gene product.	4.33e+03
682	4	20.0	237 39	W85440	A33/212 single-chain	4.33e+03	4	20.0	349 26	W14637	N. meningitidis HTR T	4.33e+03
683	4	20.0	239 39	W87763	Rice lysine ketogluta	4.33e+03	4	20.0	351 38	W86276	Signal transduction p	4.33e+03
684	4	20.0	241 39	W85441	A33/218 single-chain	4.33e+03	4	20.0	351 4	R22318	Carcino embryonic ant	4.33e+03
685	4	20.0	243 39	W82283	Tumour surface protei	4.33e+03	4	20.0	357 38	W85465	Secreted protein enco	4.33e+03
686	4	20.0	244 32	W34944	Mouse neurogenin 1 pr	4.33e+03	4	20.0	359 39	W73544	Mouse IkappaB-beta pr	4.33e+03
687	4	20.0	246 38	W83020	Osteoclastogenesis in	4.33e+03	4	20.0	365 39	W89242	Rhodococcus sp. strai	4.33e+03
688	4	20.0	246 5	R27245	Sequence of the AA se	4.33e+03	4	20.0	366 38	W82638	Erlichia sp. HOG-7 p	4.33e+03
689	4	20.0	249 39	W81550	Natural subtilisin rib	4.33e+03	4	20.0	366 39	W89228	Tumour necrosis facto	4.33e+03
690	4	20.0	254 39	W82285	Cellular killer cell p	4.33e+03	4	20.0	366 38	W86259	Maize OTE enzyme.	4.33e+03
691	4	20.0	265 28	W41152	RBE1 transcription co	4.33e+03	4	20.0	370 39	W95181	Human G-protein coupl	4.33e+03
692	4	20.0	268 39	W95358	Bovine interleukin (I	4.33e+03	4	20.0	372 39	W89449	A gida2 polypeptide f	4.33e+03
693	4	20.0	269 24	W18706	Mutant BLAP F44.	4.33e+03	4	20.0	374 39	W89438	Human G-protein coupl	4.33e+03
694	4	20.0	269 24	W18707	Mutant BLAP F43.	4.33e+03	4	20.0	375 11	R57021	Human A transferase (	4.33e+03
695	4	20.0	269 9	R46382	Subtilisin 309 Serine	4.33e+03	4	20.0	376 38	W85034	Beta subunit of DNA P	4.33e+03
696	4	20.0	269 9	R46383	PB92 serine protease,	4.33e+03	4	20.0	380 38	W88726	Secreted protein enco	4.33e+03
697	4	20.0	269 6	R30128	BLAP (S139A).	4.33e+03	4	20.0	380 6	R34613	High alkaline proteas	4.33e+03
698	4	20.0	269 15	R80415	BLAP (R99X), S154X, I2	4.33e+03	4	20.0	381 39	R13272	Polyhedrin-soluble IA	4.33e+03
699	4	20.0	269 14	R87498	Subtilisin N76D/S103A	4.33e+03	4	20.0	381 3	W86312	Kidney injury associa	4.33e+03
700	4	20.0	270 38	W88784	Polypeptide fragment	4.33e+03	4	20.0	384 39	W86318	Kidney injury associa	4.33e+03
701	4	20.0	270 2	R10568	S20 mutant of the sub	4.33e+03	4	20.0	385 4	R23793	Stearoyl-ACP-desatur	4.33e+03
702	4	20.0	271 39	W82002	Human foetal brain se	4.33e+03	4	20.0	386 39	W87884	Protein encoded by th	4.33e+03
703	4	20.0	273 39	W89748	Staphylococcus aureus	4.33e+03	4	20.0	393 4	R20802	Alpha-1-antitrypsin f	4.33e+03
704	4	20.0	274 1	P80848	Sequence of subtilisi	4.33e+03	4	20.0	396 39	W82710	Soil derived peptide	4.33e+03
705	4	20.0	274 21	W15437	Subtilisin Carlsberg	4.33e+03	4	20.0	396 39	W89680	Human osteogenic prot	4.33e+03
706	4	20.0	274 21	W15436	Subtilisin Carlsberg	4.33e+03	4	20.0	396 28	W34218	Streptomyces hydroxyl	4.33e+03
707	4	20.0	274 21	W15435	Subtilisin Carlsberg	4.33e+03	4	20.0	397 39	W89227	Tumour necrosis facto	4.33e+03
708	4	20.0	279 38	W86148	Modified dihydroptera	4.33e+03	4	20.0	398 39	W87566	A cyclin related prot	4.33e+03
709	4	20.0	282 39	W67850	Human secreted protei	4.33e+03	4	20.0	399 38	W86237	Human VEGF-C full len	4.33e+03
710	4	20.0	283 39	W73583	IDX-1 protein sequenc	4.33e+03	4	20.0	399 39	W89691	Murine osteogenic pro	4.33e+03
711	4	20.0	288 31	W57211	Human congenital hear	4.33e+03	4	20.0	399 25	W23144	Bone morphogenic prot	4.33e+03
712	4	20.0	290 39	W83329	Human mindin-related	4.33e+03	4	20.0	400 39	W73654	M. tuberculosis antig	4.33e+03
713	4	20.0	293 1	P80845	Sequence of Iritirach	4.33e+03	4	20.0	400 39	W73764	M. tuberculosis antig	4.33e+03
714	4	20.0	294 38	W830013	Expressed antigen for	4.33e+03	4	20.0	402 39	W89692	M. tuberculosis antig	4.33e+03
715	4	20.0	296 39	W87887	Protein encoded by th	4.33e+03	4	20.0	402 39	W95456	Human osteogenic prot	4.33e+03
716	4	20.0	298 37	W69303	Humanised anti-IL-8 6	4.33e+03	4	20.0	402 16	R83102	Alpha-1-antichymotryp	4.33e+03
717	4	20.0	298 6	R29509	BDV structural glycop	4.33e+03	4	20.0	402 16	R83101	Wild-type alpha-1-ant	4.33e+03
718	4	20.0	299 39	W89765	Staphylococcus aureus	4.33e+03	4	20.0	403 39	W82682	Streptomyces sp. type	4.33e+03
719	4	20.0	299 39	W80985	B cell receptor assoc	4.33e+03	4	20.0	403 18	W00084	Human Ras-related pro	4.33e+03
720	4	20.0	299 39	W80984	Amino acid sequence o	4.33e+03	4	20.0	405 39	W86329	Kidney injury associa	4.33e+03
721	4	20.0	301 39	W87740	Corynebacterium dihyd	4.33e+03	4	20.0	407 39	W95400	M. tuberculosis Rp-fa	4.33e+03
722	4	20.0	303 39	W73636	S. pneumoniae lacC pr	4.33e+03	4	20.0	411 39	W83321	Human native prouroki	4.33e+03
723	4	20.0	304 10	R53349	Wheat leaf ADP-glucos	4.33e+03	4	20.0	411 21	W13634	Human native prouroki	4.33e+03
724	4	20.0	306 39	W73641	Mouse B7-2 antigen.	4.33e+03	4	20.0	412 2	R10334	Recombinant singlech	4.33e+03
725	4	20.0	307 7	R37799	Human NGF.	4.33e+03	4	20.0	414 26	W24248	Aquifex aspartate tra	4.33e+03
726	4	20.0	309 39	W83322	Single chain Apo-2 an	4.33e+03	4	20.0	415 28	W19005	Feline herpes virus t	4.33e+03
727	4	20.0	310 39	W95197	Human syndecan-1.	4.33e+03	4	20.0	417 39	W95538	Death domain containi	4.33e+03
728	4	20.0	310 39	W83324	Single chain Apo-2 an	4.33e+03	4	20.0	417 39	W89226	Tumour necrosis facto	4.33e+03
729	4	20.0	311 39	W89229	Tumour necrosis facto	4.33e+03	4	20.0	417 26	W31517	Death domain containi	4.33e+03
730	4	20.0	312 16	R73344	Human Indian hedgehog	4.33e+03	4	20.0	419 38	W86203	Human vascular endoth	4.33e+03
731	4	20.0	312 39	W83323	Single chain Apo-2 an	4.33e+03	4	20.0	420 38	W89224	Tumour necrosis facto	4.33e+03
732	4	20.0	317 38	W88758	Polypeptide fragment	4.33e+03	4	20.0	421 38	W68541	Amino acid sequence o	4.33e+03
733	4	20.0	317 38	W83018	Osteoclastogenesis in	4.33e+03	4	20.0	421 39	W86326	Kidney injury associa	4.33e+03
734	4	20.0	317 38	W83195	Human osteoprotegerin	4.33e+03	4	20.0	423 39	W73767	M. tuberculosis antig	4.33e+03
735	4	20.0	317 38	W73507	Human ATG-1709 protei	4.33e+03	4	20.0	423 39	W73657	M. tuberculosis antig	4.33e+03
736	4	20.0	317 38	W88667	Secreted protein enco	4.33e+03	4	20.0	424 24	W24579	Inhibitor resistant m	4.33e+03
737	4	20.0	319 6	R32193	Ira-2 protein.	4.33e+03	4	20.0	428 39	W95537	Death domain containi	4.33e+03
738	4	20.0	321 39	W83375	Streptococcus pneumon	4.33e+03	4	20.0	434 38	W81458	Amino acid sequence o	4.33e+03
739	4	20.0	324 39	W83390	Caenorhabditis elegans	4.33e+03	4	20.0	435 39	W89448	A gida2 polypeptide s	4.33e+03
740	4	20.0	329 32	W60041	H. plurivialis beta-car	4.33e+03	4	20.0	441 38	W73359	Human PAF-AH protein	4.33e+03
741	4	20.0	331 39	W83328	Human mindin polypept	4.33e+03	4	20.0	441 6	R31955	Sequence encoded by g	4.33e+03
742	4	20.0	331 39	W89491	Hepatitis G virus ant	4.33e+03	4	20.0	442 38	W86283	L. infantum cysteine	4.33e+03
743	4	20.0	331 39	W89451	A gida2 polypeptide f	4.33e+03	4	20.0	444 39	W89450	A gida2 polypeptide s	4.33e+03
744	4	20.0	337 39	W89489	Hepatitis G virus ant	4.33e+03	4	20.0	444 38	W73361	Canine PAF-AH protein	4.33e+03
745	4	20.0	338 3	R21420	Streptomyces clavulig	4.33e+03	4	20.0	446 39	W83374	Streptococcus pneumon	4.33e+03
746	4	20.0	339 39	W86325	Kidney injury associa	4.33e+03	4	20.0	448 37	W82551	Human LIR-pbm2 protei	4.33e+03

820	4	20.0	448 30	W23967	Human presenilin-2.	4.33e+03	893	631 37	W82552	Human LIR-pbm17 prote	4.33e+03
821	4	20.0	449 22	W08608	Chimeric MOMLV and Ty	4.33e+03	894	632 4	R23167	Mutant thermostable D	4.33e+03
822	4	20.0	450 38	W86335	Kidney injury associa	4.33e+03	895	635 38	W75410	Fusion protein PD30 c	4.33e+03
823	4	20.0	456 38	W82657	Ehrlichia sp. extende	4.33e+03	896	635 38	W75412	Fusion protein PD33 c	4.33e+03
824	4	20.0	457 39	W85504	Mortierella alpina de	4.33e+03	897	635 38	W85034	Green fluorescent pro	4.33e+03
825	4	20.0	457 38	W85121	A delta-6 desaturase	4.33e+03	898	635 38	W75409	Fusion protein PD31 c	4.33e+03
826	4	20.0	460 39	W84181	A GDNFR-alpha-related	4.33e+03	899	638 38	W86014	R. rubrum poly-beta-h	4.33e+03
827	4	20.0	460 39	R13223	Human protein C zymog	4.33e+03	900	643 1	P91462	67-kD protein toxin.	4.33e+03
828	4	20.0	461 39	W89558	Triticum sp. cysteine	4.33e+03	901	649 38	W81977	Ehrlichia sp. B3 prot	4.33e+03
829	4	20.0	462 39	W88501	Human stomach carcino	4.33e+03	902	651 39	W86313	Kidney injury associa	4.33e+03
830	4	20.0	462 38	W83142	Chimeric receptor con	4.33e+03	903	651 39	W53875	Gravlin polypeptide fr	4.33e+03
831	4	20.0	464 39	W83041	Anti-Pas MAB HcE7A he	4.33e+03	904	652 36	W80620	S. pneumoniae putativ	4.33e+03
832	4	20.0	464 8	R42336	Human antithrombin II	4.33e+03	905	664 3	P60229	Alcohol-oxidase.	4.33e+03
833	4	20.0	473 21	W14005	Human SHC protein.	4.33e+03	906	667 3	R11974	Tobacco SURA-C3 mutan	4.33e+03
834	4	20.0	477 5	R25462	gigaA.	4.33e+03	907	674 7	R34130	Truncated PVX replica	4.33e+03
835	4	20.0	482 39	W87766	Arabidopsis saccharop	4.33e+03	908	677 38	W75411	Fusion protein PD34 c	4.33e+03
836	4	20.0	485 17	R95961	Leuconostoc mesentero	4.33e+03	909	686 4	P93711	Sequence of pseudorab	4.33e+03
837	4	20.0	488 17	R95962	Leuconostoc citreum g	4.33e+03	910	686 39	W86162	Insect GABA receptor	4.33e+03
838	4	20.0	492 25	W26609	Human agrin.	4.33e+03	911	693 1	P94856	Expression plasmid pu	4.33e+03
839	4	20.0	496 39	W89586	Aspergillus oryzae am	4.33e+03	912	698 39	W84893	4D5 Fab molecule expr	4.33e+03
840	4	20.0	496 39	W89597	Aspergillus oryzae am	4.33e+03	913	700 32	W52822	Human Nono/TFE3 fusio	4.33e+03
841	4	20.0	496 39	W87772	Human serum glucocort	4.33e+03	914	702 38	W83137	CEA protein.	4.33e+03
842	4	20.0	496 39	W89614	Aspergillus oryzae di	4.33e+03	915	716 39	W73000	Drosophila bHLH-PAS/J	4.33e+03
843	4	20.0	497 3	P60456	Sequence of aldehyde	4.33e+03	916	722 39	W86308	Kidney injury associa	4.33e+03
844	4	20.0	499 38	W67010	HCV non-structural pr	4.33e+03	917	729 39	W72999	Drosophila bHLH-PAS/I	4.33e+03
845	4	20.0	502 38	W86249	Mouse BMP receptor ki	4.33e+03	918	740 39	W95398	M. tuberculosis catal	4.33e+03
846	4	20.0	504 38	W73500	Trabecular meshwork i	4.33e+03	919	740 39	W95399	M. tuberculosis catal	4.33e+03
847	4	20.0	504 38	W80513	Hepatitis C virus ant	4.33e+03	920	748 39	W89272	Granulocytic Ehrlich	4.33e+03
848	4	20.0	504 38	W70825	Hepatitis C virus ant	4.33e+03	921	750 39	W82728	Adenovirus PNP194.SS	4.33e+03
849	4	20.0	504 39	W89391	Human trabecular mesh	4.33e+03	922	761 39	W70899	Protein encoded by tu	4.33e+03
850	4	20.0	506 39	W87626	Dirofilaria immitis a	4.33e+03	923	765 39	W89561	Recombinant heat-resi	4.33e+03
851	4	20.0	506 39	W83126	PT1K48 Lysine speci	4.33e+03	924	767 2	R07599	Sequence deduced Fl a	4.33e+03
852	4	20.0	508 2	R06512	EPO receptor.	4.33e+03	925	771 39	W89589	Aspergillus oryzae di	4.33e+03
853	4	20.0	512 39	W87797	Protease encoded by c	4.33e+03	926	774 38	W87548	A Tfu DNA polymerase	4.33e+03
854	4	20.0	514 38	W81976	Ehrlichia sp. W20.2 p	4.33e+03	927	774 30	W48459	Mutant KOD DNA polyme	4.33e+03
855	4	20.0	522 39	W81571	Mus dunni endogenous	4.33e+03	928	774 30	W48451	Mutant KOD DNA polyme	4.33e+03
856	4	20.0	526 38	W83138	FL-CEA protein.	4.33e+03	929	774 30	W48460	Mutant KOD DNA polyme	4.33e+03
857	4	20.0	529 38	W82586	Human Patched-2 prote	4.33e+03	930	796 38	W85598	Cadherin-11.	4.33e+03
858	4	20.0	529 38	W85055	Human protein designa	4.33e+03	931	797 38	W85042	PKB-green fluorescent	4.33e+03
859	4	20.0	529 38	W80498	A protein designat	4.33e+03	932	797 38	W85043	NFKappaB p65 subunit-	4.33e+03
860	4	20.0	530 38	W82656	Ehrlichia sp. extende	4.33e+03	933	822 38	W73486	Cellulose phosphoryl	4.33e+03
861	4	20.0	532 38	W83141	Chimeric receptor con	4.33e+03	934	822 5	R26337	N-sam.	4.33e+03
862	4	20.0	532 38	W86248	Mouse BMP receptor ki	4.33e+03	935	847 39	W85603	Hexosaminidase enzyme	4.33e+03
863	4	20.0	538 38	W88628	Secreted protein enco	4.33e+03	936	849 38	W82397	Human GBP protein #3.	4.33e+03
864	4	20.0	539 39	W87799	Aminopeptidase SlpD e	4.33e+03	937	865 4	R20670	Lipoxygenase.	4.33e+03
865	4	20.0	540 39	W89774	Staphylococcus aureus	4.33e+03	938	871 36	W72223	HSV-2 strain SB5 Cont	4.33e+03
866	4	20.0	540 39	W89794	Staphylococcus aureus	4.33e+03	939	875 1	P94854	C-terminal prepro-C-t	4.33e+03
867	4	20.0	543 39	W86149	Fat cell differentiat	4.33e+03	940	877 39	W90082	Porcine G-protein reg	4.33e+03
868	4	20.0	546 38	W81975	Ehrlichia sp. W20.1 p	4.33e+03	941	880 39	W90088	Human G-protein regul	4.33e+03
869	4	20.0	547 38	W81440	Human intercellular a	4.33e+03	942	908 38	W85587	Human N-methyl-D-aspa	4.33e+03
870	4	20.0	548 39	W89189	Alternatively spliced	4.33e+03	943	911 30	W54124	H. influenzae strain	4.33e+03
871	4	20.0	550 39	W81351	Human guanine nucleot	4.33e+03	944	916 38	W85023	Amino acid sequence o	4.33e+03
872	4	20.0	551 39	W73480	Grapevine leafroll vi	4.33e+03	945	917 38	W73512	Rat ICAM-4 protein se	4.33e+03
873	4	20.0	555 39	W89799	Staphylococcus aureus	4.33e+03	946	929 38	W87501	N-methyl-D-aspartate	4.33e+03
874	4	20.0	557 39	W81996	A. thaliana trehalase	4.33e+03	947	936 39	W89601	Staphylococcus aureus	4.33e+03
875	4	20.0	575 38	W83140	Chimeric receptor con	4.33e+03	948	947 39	W82497	Human NIK protein.	4.33e+03
876	4	20.0	576 39	W88523	Eat-4 protein amino a	4.33e+03	949	947 38	W81561	Human NF-kB-inducing	4.33e+03
877	4	20.0	580 39	W81349	Human guanine nucleot	4.33e+03	950	947 38	W81564	Human NF-kB-inducing	4.33e+03
878	4	20.0	581 39	W81993	35S CaMV anti-sense t	4.33e+03	951	948 38	W83148	Rat receptor tyrosine	4.33e+03
879	4	20.0	584 3	R14670	Truncated poly Ig-rec	4.33e+03	952	976 38	W85581	Human N-methyl-D-aspa	4.33e+03
880	4	20.0	590 39	W87628	Dirofilaria immitis a	4.33e+03	953	990 39	W90075	P. vulgaris chondroit	4.33e+03
881	4	20.0	590 39	W87629	Dirofilaria immitis a	4.33e+03	954	997 38	W85589	Human N-methyl-D-aspa	4.33e+03
882	4	20.0	590 38	W82649	Ehrlichia sp. extende	4.33e+03	955	1012 39	W95395	IBDV VP2-VP3-VP4 poly	4.33e+03
883	4	20.0	591 38	W73316	Parathyroid hormone r	4.33e+03	956	1022 39	W87761	Maize lysine ketoglut	4.33e+03
884	4	20.0	593 39	W81974	Lactococcus lactis Lt	4.33e+03	957	1027 39	W89806	Staphylococcus aureus	4.33e+03
885	4	20.0	599 39	W73479	Grapevine leafroll vi	4.33e+03	958	1038 38	W86247	Mouse BMP receptor ki	4.33e+03
886	4	20.0	600 39	W82660	Cauliflower L-galacto	4.33e+03	959	1038 38	W86246	Human BMP receptor ki	4.33e+03
887	4	20.0	606 38	W84086	Human membrane fusio	4.33e+03	960	1064 39	W87759	Arabidopsis lysine ke	4.33e+03
888	4	20.0	609 39	W82726	Adenovirus PACTSsigDE	4.33e+03	961	1080 3	P50296	Mouse epidermal growt	4.33e+03
889	4	20.0	609 39	W82727	Adenovirus P1gKN pro	4.33e+03	962	1092 39	W89257	Gracilariopsis lemane	4.33e+03
890	4	20.0	619 39	W73532	MEKK2 protein	4.33e+03	963	1102 39	W90083	Porcine G-protein reg	4.33e+03
891	4	20.0	622 39	W81570	Mus dunni endogenous	4.33e+03	964	1105 32	W44864	Human TPC2 telomere r	4.33e+03
892	4	20.0	625 38	W83200	Murine osteoclast dif	4.33e+03	965	1118 38	W82395	Human GBP protein #1.	4.33e+03

```
966 4 20.0 1155 39 W83393 Bacillus thuringiensis 4.33e+03
967 4 20.0 1167 39 W87632 CryET4 protein sequen 4.33e+03
968 4 20.0 1174 38 W67438 Human protein tyrosin 4.33e+03
969 4 20.0 1195 39 W75420 T.thermophilus nitrat 4.33e+03
970 4 20.0 1195 1 R05530 High density lipoprot 4.33e+03
971 4 20.0 1203 39 W81572 Mus dunni endogenous 4.33e+03
972 4 20.0 1229 39 W87633 CryET5 protein sequen 4.33e+03
973 4 20.0 1253 39 W88524 Adenyl cyclase-1 (acy 4.33e+03
974 4 20.0 1275 39 W82594 H. contortus PCP-A pr 4.33e+03
975 4 20.0 1287 15 R79844 Helicobacter pylori v 4.33e+03
976 4 20.0 1305 39 W86525 Adenyl cyclase type 9 4.33e+03
977 4 20.0 1306 1 R04111 Human angiotensin con 4.33e+03
978 4 20.0 1336 38 W87511 Human N-methyl-D-aspa 4.33e+03
979 4 20.0 1456 14 R70419 Rat homologue of huma 4.33e+03
980 4 20.0 1464 38 W92994 An antigen from dermo 4.33e+03
981 4 20.0 1479 28 W41118 Murine type C lectin 4.33e+03
982 4 20.0 1481 39 W89299 Human rchd528 protein 4.33e+03
983 4 20.0 1523 38 W87554 A Tfu DNA polymerase 4.33e+03
984 4 20.0 1663 27 W34619 Human C3 protein muta 4.33e+03
985 4 20.0 1663 27 W34620 Human C3 protein muta 4.33e+03
986 4 20.0 2050 38 W73499 Von Willebrand factor 4.33e+03
987 4 20.0 2100 39 W89579 Calcium permeable vol 4.33e+03
988 4 20.0 2104 39 W89578 Calcium permeable vol 4.33e+03
989 4 20.0 2105 39 W89577 Calcium permeable vol 4.33e+03
990 4 20.0 2154 39 W81639 Mouse elf-1 protein. 4.33e+03
991 4 20.0 2261 1 P90164 Peptide encoded by co 4.33e+03
992 4 20.0 2337 32 W37878 Human calcium channel 4.33e+03
993 4 20.0 2351 24 W11454 Active Factor VIII:C 4.33e+03
994 4 20.0 2351 2 P70448 Human factor VIII:C 4.33e+03
995 4 20.0 2352 24 W11357 Active Factor VIII:C 4.33e+03
996 4 20.0 2349 32 W56027 FRAP (torl) protein. 4.33e+03
997 4 20.0 2639 39 W73476 Grapevine leafroll vi 4.33e+03
998 4 20.0 3119 26 W36888 Mouse Huntington's di 4.33e+03
999 4 20.0 3729 26 W22603 Tylosone synthase OR 4.33e+03
1000 4 20.0 4303 17 R90302 Polycystic kidney dis 4.33e+03
```

## ALIGNMENTS

```
RESULT 1
ID R71833 standard; Protein; 113 AA.
AC R71833;
DF 08-NOV-1995 (first entry)
DE Alternaria alternata allergen Altall.
KW Fungal spore; allergen; Altall; allergy; ribosomal protein P2.
OS Alternaria alternata.
PN W09506122-A.
PD 02-MAR-1995.
PF 24-AUG-1994; AT0121.
PR 27-AUG-1993; AT-001726.
PA (BIOM-) BIOMAY PROD N & HANDELSGES MBH.
PI Achatz G, Breitenbach M, Ebner C, Hirschwehr R;
PI Kraft D, Lechenauer E, Oberkofler H, Prillinger H;
PI Simon B, Unger A;
DR WPI: 95-106851/14.
DR N-PSDB: Q86276.
PT Allergens derived from Alternaria alternata - their isolation by
PT reacting sera from patients with fungal extracts; useful for
PT in-vitro allergy detection.
PS Claim 7; Page 13; 26pp; German.
CC Spores of Alternaria alternata can cause allergic reactions. Various
CC Alta allergens and sequences encoding them have now been isolated.
CC The mature Altall allergen has mol. wt. 11 kD and is encoded by cDNA
CC sequence Q86276. The allergen has homology to ribosomal protein P2.
CC Potential epitopic subfragments were identified by computer analysis
CC of the amino acid sequence. See R71834-R71838 for potential B-cell
CC epitopes and R71839-R71840 for potential T-cell epitopes.
SQ Sequence 113 AA;
```

Query Match 30.0%; Score 6; DB 13; Length 113;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 73 aaasgg 78
Qy 15 AAASGG 20

RESULT 2
ID R26055 standard; Protein; 123 AA.
AC R26055;
DT 01-FEB-1993 (first entry)
DE Human galanin.
KW Fertility; pain; irritable bowel syndrome; anorexia; insulin;
KW hypersecretion; growth hormone deficiency; intestinal motility;
KW constipation; obesity; diabetes; pseudoobstruction.
OS Homo sapiens.
FH Key Location/Qualifiers
FT region 1..33 /note= "signal peptide"
FT peptide 34..64 /note= "galanin"
FT region 65..123 /note= "extension peptide"
PN W09212997-A.
PD 06-AUG-1992.
PF 16-JAN-1992; U00368.
PR 16-JAN-1991; US-641344.
PA (GPHO ) GEN HOSPITAL CORP.
PI Kaplan LM;
DR WPI: 92-284601/34.
DR N-PSDB: Q27236.
PT Human galanin and polypeptide galanin antagonist - for use in
PT birth control and for treating e.g. pain, irritable bowel
PT syndrome, diabetes etc.
PS Claim 4; Fig 1; 20pp; English.
CC Galanin is encoded as part of a precursor peptide comprising 123
CC amino acids including a signal sequence and an extension peptide
CC of 59 amino acids. This peptide or its analogues and fragments
CC esp. 2-15, 2-23, 15-30 and 21-30 of mature galanin have galanin
CC activity and may be used in therapeutic compsns. for decreasing
CC fertility in women and for treating pain, irritable bowel syndrome,
CC anorexia, insulin hypersecretion and growth hormone deficiency in
CC humans. Inhibitory peptide fragments of galanin may be used as
CC galanin antagonists which are useful for increasing fertility in
CC women, increasing intestinal motility and treating constipation
CC ileus, gastroparesis diabeticorum, chronic pseudoobstruction,
CC obesity and diabetes.
SQ Sequence 123 AA;
```

Query Match 30.0%; Score 6; DB 5; Length 123;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 111 lpaas 116
Qy 13 LPAAAS 18
```

```
RESULT 3
ID R27115 standard; Protein; 123 AA.
AC R27115;
DT 25-FEB-1993 (first entry)
DE Sequence of human preprogalanin and galanin mRNA associated
DE peptide (GMAP).
KW Galanin; neuropeptide; growth hormone stimulator; diarrhoea;
KW therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..32 /label= prepropeptide
FT protein 33..62 /label= galanin
FT peptide 65..123 /note= "Clained"
FT peptide /label= GMAP
```

```

PN WO9215681-A.
PD 17-SEP-1992.
PF 06-MAR-1992; A00097.
PR 06-MAR-1991; AU-004953.
PA (GARV- ) GARVAN INST MEDICAN RES.
PI Evans HF, Shine J;
DR WPI: 92-331727/40.
DR N-PSDB: Q28890.
PT Human galanin polypeptide - as an insulin secretion inhibitor and
PT growth hormone stimulator; also treats diarrhoea, modulates endocrine
PT pancreatic disorders, attenuates cardiac vagal function etc.
PS Disclosure: Fig 1: 49pp; English.
CC Oligonucleotides complementary to two conserved regions of pig and
CC rat preprogalanin were used in a polymerase chain reaction (PCR) to
CC specifically amplify the corresp. seq. from neuroblastoma and
CC pituitary cDNA. The two amplification oligos used (No. 1 and 2)
CC corresp. to AAs 29-37 and 105-97 of rat and pig preprogalanin
CC respectively, and flank a 230 bp region encoding galanin and the
CC N-terminus of GMAP. An additional oligonucleotide (No. 3) within
CC this region was used to probe for the correct PCR product. The
CC amplified region from both sources of DNA was subcloned and then
CC sequenced revealing identical sequences. Human preprogalanin cDNA
CC is different to that of pig, cow and rat. Consequently, human
CC galanin may have a variety of biological properties that differ
CC from porcine, rat and bovine galanin.
SQ Sequence 123 AA;

Query Match 30.0%; Score 6; DB 5; Length 123;
Best Local Similarity 100.0%; Pred. No. 9.42e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 lpaas 116
Qy 13 LPAAS 18

RESULT 4
ID P09051 standard; protein; 153 AA.
AC P09051:
DT 16-NOV-1989 (first entry)
DE Herpes Simplex virus-1 730 bp open reading frame.
KW Herpes Simplex virus-1; 730 bp open reading frame; diagnostic test kit;
KW membrane-free polypeptide; antibody; anti-antibody; detection.
OS Herpes simplex virus-1
PN AU930061-A.
PD 22-JUN-1989.
PF 17-FEB-1989; 030061.
PR 30-AUG-1983; AU-030061.
PA (GETH) Genentech Inc.
PI Berman FW; Lasky LA.
DR WPI: 89-234099/33.
PT Diagnostic test kit - contains molecularly cloned, truncated
PS Disclosure: fig 15; 88pp; English.
CC membrane-free deriv. of polypeptide and complementary antibody
CC or anti-antibody.
CC Herpes Simplex virus-1 730 bp ORF protein. It was compared with the
CC corresp. ORF of HSV-2 (see P090503). The patent claims truncated
CC membrane free derivs. of HSV genes, with antigenic determinants
CC capable of specifically binding complementary antibody (Ab), and
CC a diagnostic kit contg the deriv. and the Ab or anti-Ab.
CC for detecting HSV infection. See N0524-6,N0529-30, P90165,
CC and P90166.
SQ Sequence 153 AA;

Query Match 30.0%; Score 6; DB 1; Length 153;
Best Local Similarity 100.0%; Pred. No. 9.42e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 rlpaa 9
Qy 12 RLPAA 17

```

```

RESULT 5
ID R88652 standard; Protein; 198 AA.
AC R88652:
DT 02-SEP-1996 (first entry)
DE N.meningitidis IM2169 Tbp2 domain 1 and domain 3 deletion mutant.
DE Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;
KW passive immunisation; immunotherapy; domain 2.
KW Neisseria meningitidis (strain IM2169).
OS Synthetic.
FH Key Location/Qualifiers
FT domain 1..198
FT /label= Domain_2
FT /note= "complete hinge domain; the entire domain 1
FT and domain 3 sequences have been deleted"
PN WO9533049-A2.
PD 07-DEC-1995.
PR 30-MAY-1995; F00701.
PR 31-MAY-1994; FR-006594.
PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.
PA (TRGE ) TRANSCENE SA.
PI Jacobs E, Legrain M, Mazarin V, Lissolo L;
PI Millet MBJ;
DR WPI: 96-030562/03.
PT Polypeptide(s) for vaccination against Neisseria meningitidis group
PT B - comprising deletion mutants of transferrin receptor Tbp2
PT subunit
PS Claim 1; Page -: 114pp; French.
CC Deletion mutants derived from the N.meningitidis strain IM2169
CC transferrin Tbp2 subunit are claimed in which at least one of the
CC three domains is partially or totally deleted, provided that the
CC first and second domains are not simultaneously partially or totally
CC deleted. Derivatives of IM2169-related strains are also claimed and
CC for these polypeptides the positions of the 3 domains are defined by
CC alignment with the IM2169 sequence. The deletion mutant polypeptides
CC of the invention can generate an immune response against
CC N.meningitidis. The present sequence is that of a preferred mutant
CC in which domains 1 and 3 have been deleted in their entirety, i.e.
CC it is the sequence of domain 2 of the IM2169 Tbp2 subunit.
SQ Sequence 198 AA;

Query Match 30.0%; Score 6; DB 17; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.42e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 aaasgg 22
Qy 15 AAASGG 20

RESULT 6
ID R88660 standard; Protein; 198 AA.
AC R88660:
DT 03-SEP-1996 (first entry)
DE N.meningitidis S3032 Tbp2 domain 1 and domain 3 deletion mutant.
DE Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;
KW passive immunisation; immunotherapy; domain 2.
KW Neisseria meningitidis (strain S3032).
OS Synthetic.
FH Key Location/Qualifiers
FT domain 1..198
FT /label= Domain_2
FT /note= "complete hinge domain; the entire domain 1
FT and domain 3 sequences have been deleted"
PN WO9533049-A2.
PD 07-DEC-1995.
PR 30-MAY-1995; F00701.
PR 31-MAY-1994; FR-006594.
PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.
PA (TRGE ) TRANSCENE SA.
PI Jacobs E, Legrain M, Mazarin V, Lissolo L;
PI Millet MBJ;
DR WPI: 96-030562/03.
PT Polypeptide(s) for vaccination against Neisseria meningitidis group

```



B - comprising deletion mutants of transferrin receptor Tbp2 subunit

Claim 1: Page 70; 114pp; French.

Deletion mutants derived from the N.meningitidis strain IM2169 transferrin Tbp2 subunit are claimed in which at least one of the three domains is partially or totally deleted, provided that the first and second domains are not simultaneously partially or totally deleted. Derivatives of IM2169-related strains are also claimed and for these polypeptides the positions of the 3 domains are defined by alignment with the IM2169 sequence. The deletion mutant polypeptides of the invention can generate an immune response against N.meningitidis. The present sequence is that of a preferred mutant in which domains 1 and 3 of the IM2169-related strain S3032 have been deleted in their entirety, i.e. it is the sequence of domain 2 of the S3032 Tbp2 subunit.

Sequence 198 AA;

Query Match 30.0%; Score 6; DB 17; Length 199;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 aaasgg 24  
|||||

Qy 15 AAASGG 20

RESULT 7

ID R88661 standard; Protein; 199 AA.

AC R88661;

DE 03-SEP-1996 (first entry)

DT N.meningitidis Bz83 Tbp2 domain 1 and domain 3 deletion mutant.

KW Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;

KV passive immunisation; immunotherapy; domain 2.

OS Neisseria meningitidis (strain Bz83).

OS Synthetic.

FH Key Location/Qualifiers

FT domain 1..199

FT /label= Domain\_2

FT /note= "complete hinge domain; the entire domain 1 and domain 3 sequences have been deleted"

PN W09533049-A2.

PD 07-DEC-1995.

PF 30-MAY-1995; F00701.

PR 31-MAY-1994; FR-006594.

PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.

PA (TRGE ) TRANSGENE SA.

PI Jacobs E, Legrain M, Mazarin V, Lissolo L;

PI Millet MBJ;

DR WPI; 96-030562/03.

PT Polypeptide(s) for vaccination against Neisseria meningitidis group B - comprising deletion mutants of transferrin receptor Tbp2 subunit

PS Claim 1: Page 70; 114pp; French.

CC Deletion mutants derived from the N.meningitidis strain IM2169 transferrin Tbp2 subunit are claimed in which at least one of the three domains is partially or totally deleted, provided that the first and second domains are not simultaneously partially or totally deleted. Derivatives of IM2169-related strains are also claimed and for these polypeptides the positions of the 3 domains are defined by alignment with the IM2169 sequence. The deletion mutant polypeptides of the invention can generate an immune response against N.meningitidis. The present sequence is that of a preferred mutant in which domains 1 and 3 of the IM2169-related strain Bz83 have been deleted in their entirety, i.e. it is the sequence of domain 2 of the Bz83 Tbp2 subunit.

Sequence 199 AA;

Query Match 30.0%; Score 6; DB 17; Length 199;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 aaasgg 22  
|||||

QY 15 AAASGG 20

RESULT 8

ID R88658 standard; Protein; 211 AA.

AC R88658;

DT 03-SEP-1996 (first entry)

DE N.meningitidis M978 Tbp2 domain 1 and domain 3 deletion mutant.

KW Transferrin receptor; Tbp2 subunit; deletion mutant; vaccine;

KV passive immunisation; immunotherapy; domain 2.

OS Neisseria meningitidis (strain M978).

OS Synthetic.

FH Key Location/Qualifiers

FT domain 1..211

FT /label= Domain\_2

FT /note= "complete hinge domain; the entire domain 1 and domain 3 sequences have been deleted"

PN W09533049-A2.

PD 07-DEC-1995.

PF 30-MAY-1995; F00701.

PR 31-MAY-1994; FR-006594.

PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.

PA (TRGE ) TRANSGENE SA.

PI Jacobs E, Legrain M, Mazarin V, Lissolo L;

PI Millet MBJ;

DR WPI; 96-030562/03.

PT Polypeptide(s) for vaccination against Neisseria meningitidis group B - comprising deletion mutants of transferrin receptor Tbp2 subunit

PS Claim 1: Page 67-68; 114pp; French.

CC Deletion mutants derived from the N.meningitidis strain IM2169 transferrin Tbp2 subunit are claimed in which at least one of the three domains is partially or totally deleted, provided that the first and second domains are not simultaneously partially or totally deleted. Derivatives of IM2169-related strains are also claimed and for these polypeptides the positions of the 3 domains are defined by alignment with the IM2169 sequence. The deletion mutant polypeptides of the invention can generate an immune response against N.meningitidis. The present sequence is that of a preferred mutant in which domains 1 and 3 of the IM2169-related strain M978 have been deleted in their entirety, i.e. it is the sequence of domain 2 of the M978 Tbp2 subunit.

Sequence 211 AA;

Query Match 30.0%; Score 6; DB 17; Length 211;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 aaasgg 16  
|||||

QY 15 AAASGG 20

RESULT 9

ID R23599 standard; Protein; 322 AA.

AC R23599;

DT 20-OCT-1992 (first entry)

DE Recombinant hematopoietic molecule 4.

KW IL-3; EPO; haematopoiesis.

OS Homo sapiens.

PN W09206116-A.

PD 16-APR-1992.

PF 26-SEP-1991; U07053.

PR 28-SEP-1990; US-589958.

PA (ORTH ) ORTHO PHARM CORP.

PI Rosen JI;

DR WPI; 92-150819/18.

PT Recombinant haematopoietic molecules useful in treating anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF and has early and later myeloid differentiation activity

PS Disclosure; Page 39; 82pp; English.

CC This protein sequence given comprises the entire amino acid sequence of a recombinant haematopoietic molecule, with the amino portion



CC comprising EPO and the carboxyl portion comprising IL-3. (Specific  
 CC sequences for these portions are given in R23591 and R23593.)  
 CC Within the scope of the invention hybrid molecules were produced which  
 CC contain at least a portion of an early MDF and at least a portion of a  
 CC late MDF covalently linked. These compounds can be used to promote  
 CC hematopoiesis in a patient.  
 CC The bonding of the early and late factors allows a very high conc. of  
 CC late MDF at the surface of a cell which the early MDF is bound. It  
 CC also allows the early MDF to act more specifically to stimulate only  
 CC the desired lineage, thus reducing undesirable effects. These  
 CC compounds are useful for treating anaemias of various origins eg. renal  
 CC failure and AIDS. It is easier to produce and administer one  
 CC recombinant molecule rather than two separate molecules.  
 SQ Sequence 322 AA;

Query Match 30.0%; Score 6; DB 4; Length 322;  
 Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 167 aaasgg 172  
 |||||  
 Qy 15 AAASGG 20

RESULT 10  
 ID R23079 standard; Protein: 349 AA.

AC R23079;  
 DT 20-OCT-1992 (first entry)  
 DE Epo:IL-3 Flex, recombinant hematopoietic molecule.  
 KW Early MDF; late MDF; haematopoiesis; Epo; IL-3; linker;  
 KW growth factor.  
 OS Homo sapiens.

FH Key Location/Qualifiers  
 FT peptide 1..27  
 FT /label= sig\_peptide  
 FT protein 28..349  
 FT /label= mat\_protein

FN W09206116-A.

PD 16-APR-1992.

PF 26-SEP-1991; U07053.

PR 28-SEP-1990; US-589958.

PA (ORTH ) ORTHO PHARM CORP.

PI Rosen JI;

DR WPI: 92-150819/19.

DR N-PSDB: Q24285.

PT Recombinant haematopoietic molecules useful in treating  
 PT anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF  
 PT and has early and later myeloid differentiation activity

PS Disclosure: Page 51; 82pp; English.

CC The amino acid sequence given is an Epo:IL-3 hybrid growth factor  
 CC derived from a construction formed by ligating the native Epo signal  
 CC sequence and various synthetic oligonucleotides corresponding to Epo  
 CC and IL-3 gene sequences. This molecule is comparable to the sequence  
 CC given in R23076 and contains a flexible linker molecule. This hybrid  
 CC growth factor is a haematopoietic molecule which contains at least a  
 CC portion of an early MDF and at least a portion of a late MDF  
 CC covalently linked. This compound can be used to promote hematopoiesis  
 CC in a patient.

CC The bonding of the early and late factors allows a very high conc. of  
 CC late MDF at the surface of a cell which the early MDF is bound. It  
 CC also allows the early MDF to act more specifically to stimulate only  
 CC the desired lineage, thus reducing undesirable effects. These  
 CC compounds are useful for treating anaemias of various origins eg. renal  
 CC failure and AIDS. It is easier to produce and administer one  
 CC recombinant molecule rather than two separate molecules.  
 SQ Sequence 349 AA;

Query Match 30.0%; Score 6; DB 4; Length 349;  
 Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 194 aaasgg 199  
 |||||

Qy 15 AAASGG 20

RESULT 11

ID R20178 standard; Protein: 353 AA.

AC R20178;

DT 08-APR-1992 (first entry)

DE P. glumae Pgl lipase stabilisation/translocation protein.

KW lipase-specific stabilisation/translocation function; detergent;

KW cleaning system; ORF2.

OS pseudomonas glumae.

PN EP-464922-A.

PD 08-JAN-1992.

PF 26-JUN-1991; 201627.

PR 06-JUL-1990; EP-307440.

PR 09-JUL-1990; WO-E01138.

PR 17-OCT-1990; EP-202772.

PA (UNIL ) UNILEVER NV.

PI Bos JW, Frenken LCJ, Verrips CT, Visser C;

DR WPI: 92-010138/02.

DR N-PSDB: Q20372.

PT Microorganism for use in detergents - contains lipase gene and a

PT translocation or stabilisation gene, used to produce an efficient

PT stable lipase

PS Disclosure: Fig 3; 46pp; English.

CC A P. glumae Pgl genomic library was constructed in cosmid vector c2RB.

CC The lipA gene was isolated by screening with probes based on the

CC N-terminal amino acid sequence of the P. glumae lipase. Immediately

CC downstream of the lipA gene was found a second open reading frame,

CC designated ORF2. It is ORF2 which encodes the lipase-specific

CC stabilisation/translocation activity. See also Q20373 and Q20464.

CC Sequence 353 AA;

Query Match 30.0%; Score 6; DB 3; Length 353;

Best Local Similarity 100.0%; Pred. No. 9.42e+01;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 62 paaasg 67

|||||

Qy 14 PAAASG 19

RESULT 12

ID P80277 standard; protein: 374 AA.

AC P80277;

DT 21-MAR-1991 (first entry)

DE Sequence encoded by 3' end of phenylalanine ammonia lyase

DE (PAL) CDNA derived from Rhodosporidium toruloides.

KW Phenylketonuria therapy; L-phenylalanine therapy; enzyme: ss.

OS Rhodosporidium toruloides.

PN W08802024-A.

PD 24-MAR-1988.

PF 08-SEP-1987; G006528

PR 08-SEP-1986; GB-021626.

PA (PUBL-) Public Health Lab (ANSO/).

PI Anson J, Gilbert H, Oram J, Minton NP;

DR WPI: 88-091814/13.

DR N-PSDB: N80457.

PT Gene encoding phenylalanine ammonia lyase - obt'd. from a strain of

PT R. toruloides by excising non-coding introns.

PS Example; Fig 5; 43pp; English..

CC Figure 3 (N80456) shows a portion of the genetic DNA sequence of R.

CC toruloides which encodes PAL. The gene of the invention preferably

CC consists of a DNA sequence identical to, related to, derived from or

CC complementary to the sequence of codons from the start codon to the

CC stop codon in Fig 3, from which the six intron IVS1 to IVS6 have

CC been deleted (see N82309). The gene from which introns have been

CC deleted is prep'd. by joining overlapping cDNA clones pPAL1 (Fig 5,

CC N80457) and pPAL2 (Fig 6, N80458). As the introns present in the

CC natural PAL gene act as a barrier to the expression of PAL in

CC organisms other than R. toruloides, the gene can be used to produce

CC PAL in a wide range of procaryotic and eukaryotic hosts.

CC Sequence 374 AA;

SQ

Query Match 30.0%; Score 6; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 238 tlakrl 243  
| | | | |  
Qy 8 TLAKRL 13

RESULT 13  
ID R31036 standard; Protein; 426 AA.  
AC R31036;  
DE 26-MAY-1993 (first entry)  
KW C. acidovorans stereospecific hydrolase.  
KW S-(+)-2,2-dimethylcyclopropane carboxamide; 2,2-DMCPCA; prodn.;  
KW cilastatin; antibiotic therapy; penem; carabapenem.  
OS Comamonas acidovorans A:18.  
FH Key Location/Qualifiers  
FT peptide 1..7  
FT peptide /note= "N-terminal peptide used to obtain Q35116"  
FT peptide 13..18  
FT peptide /note= "peptide used to obtain Q35117"

FN EP-524604-A.  
FD 27-JAN-1993.  
PF 21-JUL-1992; 112446.  
PR 26-JUL-1991; CH-002247.  
PA (LONZ ) LONZA AG.  
PI Birch OM, Bohlen E, Robins K, Zimmermann T;  
DR WPI: 93-028686/04.  
DR N-PSDB: Q38949.  
PT Pure S-(+)-2,2-di:methyl:cyclopropane carboxamide prodn. by  
PT microorganisms(s) - with a stereospecific hydrolase isolated from  
PT microorganism, useful for inhibition of renal dehydropeptidase  
PT via Cilastatin  
PS Disclosure; Fig 3; 19pp; German.  
CC The sequence is that of Comamonas acidovorans A:18 stereospecific  
CC hydrolase which can be used in the prodn. of optically pure S-(+)-2,2-  
CC dimethylcyclo propane carboxamide (2,2-DMCPCA); the starting  
CC material for the prodn. of cilastatin. Cilastatin inhibits e.g. renal  
CC dehydropeptidase (RDP) and in therapy it is administered with the  
CC antibiotic penem or carabapenem to prevent their inactivation by RDP.  
SQ Sequence 426 AA;

Query Match 30.0%; Score 6; DB 6; Length 426;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 350 rlpaaa 355  
| | | | |  
Qy 12 RLPAAA 17

RESULT 14  
ID W36485 standard; Protein; 459 AA.  
AC W36485;  
DE 27-APR-1998 (first entry)  
KW Mouse TUB Form I.  
KW TULP; TUB Form I; sensory neuron; neurosensory defect;  
KW cochlear degeneration; hearing loss; deafness; retinal dystrophy;  
KW retinitis pigmentosa; combined rod cone dystrophy; obesity; mouse;  
KW animal model; transgenic animal; therapy; diagnosis.  
OS Mus musculus.  
PN W09738004-A1.  
PD 16-OCT-1997.  
PF 10-APR-1997; U05903.  
PR 17-SEP-1996; US-714991.  
PR 10-APR-1996; US-630592.  
PR 22-AUG-1996; US-701380.  
PR 04-SEP-1996; US-706292.  
PA (JACK-) JACKSON LAB.  
PA (SEQU-) SEQUANA THERAPEUTICS INC.  
PI Naggett J, Nishina P, Noben-Trauth K, North M;  
DR WPI: 97-512642/47.  
DR N-PSDB; T96637.  
PT Mammalian TULP protein - used for detecting pre-disposition to  
PT neuro-sensory defects  
PS Disclosure; Page 45-47; 89pp; English.  
CC This sequence comprises mouse TUB Form II, a member of the mammalian  
CC TULP gene family associated with various defects in sensory neurons  
CC such as cochlear defects, retinitis pigmentosa and combined  
CC rod-cone dystrophy. Its amino acid sequence was deduced from a  
CC cDNA clone (see T96637) obtained from a testis cDNA library. The  
CC protein is expressed most strongly in the eye, brain and testis.  
CC TUB mutation is also associated with a genetic predisposition to  
CC adult onset obesity. Sequences are provided for human and mouse  
CC TULP proteins (see W36485-98). These proteins are useful as  
CC immunogens to raise antibodies that specifically identify TULP  
CC expressing cells, in drug screening assays directed at neurosensory  
CC defects, and for therapeutic purposes.  
SQ Sequence 459 AA;

Query Match 30.0%; Score 6; DB 27; Length 505;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 113 aaasgg 118  
| | | | |

DR WPI: 97-512642/47.  
DR N-PSDB; T96636.  
PT Mammalian TULP protein - used for detecting pre-disposition to  
PT neuro-sensory defects  
PS Disclosure; Page 44-45; 89pp; English.  
CC This sequence comprises mouse TUB Form I, a member of the mammalian  
CC TULP gene family associated with various defects in sensory neurons  
CC such as cochlear defects, retinitis pigmentosa and combined  
CC rod-cone dystrophy. Its amino acid sequence was deduced from a  
CC cDNA clone (see T96636) obtained from a testis cDNA library. The  
CC protein is expressed most strongly in the eye, brain and testis.  
CC TUB mutation is also associated with a genetic predisposition to  
CC adult onset obesity. Sequences are provided for human and mouse  
CC TULP proteins (see W36485-98). These proteins are useful as  
CC immunogens to raise antibodies that specifically identify TULP  
CC expressing cells, in drug screening assays directed at neurosensory  
CC defects, and for therapeutic purposes.  
SQ Sequence 459 AA;

Query Match 30.0%; Score 6; DB 27; Length 459;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 aaasgg 72  
| | | | |  
Qy 15 AAASGG 20

RESULT 15  
ID W36486 standard; Protein; 505 AA.  
AC W36486;  
DE 27-APR-1998 (first entry)  
DE Mouse TUB Form II.  
KW TULP; TUB Form II; sensory neuron; neurosensory defect;  
KW cochlear degeneration; hearing loss; deafness; retinal dystrophy;  
KW retinitis pigmentosa; combined rod cone dystrophy; obesity; mouse;  
KW animal model; transgenic animal; therapy; diagnosis.  
OS Mus musculus.  
PN W09738004-A1.  
PD 16-OCT-1997.  
PF 10-APR-1997; U05903.  
PR 17-SEP-1996; US-714991.  
PR 10-APR-1996; US-630592.  
PR 22-AUG-1996; US-701380.  
PR 04-SEP-1996; US-706292.  
PA (JACK-) JACKSON LAB.  
PA (SEQU-) SEQUANA THERAPEUTICS INC.  
PI Naggett J, Nishina P, Noben-Trauth K, North M;  
DR WPI: 97-512642/47.  
DR N-PSDB; T96637.  
PT Mammalian TULP protein - used for detecting pre-disposition to  
PT neuro-sensory defects  
PS Disclosure; Page 45-47; 89pp; English.  
CC This sequence comprises mouse TUB Form II, a member of the mammalian  
CC TULP gene family associated with various defects in sensory neurons  
CC such as cochlear defects, retinitis pigmentosa and combined  
CC rod-cone dystrophy. Its amino acid sequence was deduced from a  
CC cDNA clone (see T96637) obtained from a testis cDNA library. The  
CC protein is expressed most strongly in the eye, brain and testis.  
CC TUB mutation is also associated with a genetic predisposition to  
CC adult onset obesity. Sequences are provided for human and mouse  
CC TULP proteins (see W36485-98). These proteins are useful as  
CC immunogens to raise antibodies that specifically identify TULP  
CC expressing cells, in drug screening assays directed at neurosensory  
CC defects, and for therapeutic purposes.  
SQ Sequence 505 AA;

Query Match 30.0%; Score 6; DB 27; Length 505;  
Best Local Similarity 100.0%; Pred. No. 9.42e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 113 aaasgg 118  
| | | | |

QY 15 AAASG 20

Search completed: Sat Aug 28 14:50:52 1999  
Job time : 70 secs.

**This Page Blank (uspto)**

SUMMARIES									
result No.	Score	Query Match	Length	DB	ID	Description	pred. No.		
1	8	40.0	1320	2	S66279	proline dehydrogenase	1.10e-02		
2	7	35.0	153	2	S68406	vesicular integral m	5.09e-01		
3	6	30.0	35	2	S61547	transferrin binding p	1.78e-01		
4	6	30.0	107	2	S59919	acidic ribosomal prot	1.78e-01		
5	6	30.0	107	2	S59913	acidic ribosomal prot	1.78e-01		
6	6	30.0	113	2	S43109	acidic ribosomal prot	1.78e-01		
7	6	30.0	120	2	T02037	acidic ribosomal prot	1.78e-01		
8	6	30.0	123	1	RHHUN	galactin precursor - h	1.78e-01		
9	6	30.0	157	2	A31105	fimbrial protein prec	1.78e-01		
10	6	30.0	168	2	S22348	H <sup>+</sup> -transporting ATP s	1.78e-01		
11	6	30.0	172	1	WMBEF5	UL45 protein - human	1.78e-01		
12	6	30.0	196	2	S06613	chorion protein s19 -	1.78e-01		
13	6	30.0	214	2	G70971	hypothetical protein	1.78e-01		
14	6	30.0	217	2	S66044	deoxyurine kinase su	1.78e-01		
15	6	30.0	245	2	T03413	probable 1-aminocyclo	1.78e-01		
16	6	30.0	247	2	C69954	hypothetical protein	1.78e-01		
17	6	30.0	258	2	S33355	myosin heavy chain an	1.78e-01		
18	6	30.0	259	2	A47237	G-box-binding factor	1.78e-01		
19	6	30.0	261	2	S59629	glutathione transfera	1.78e-01		
20	6	30.0	262	2	S57276	14-3-3 protein homolo	1.78e-01		
21	6	30.0	265	2	S24993	chlorophyll a/b-bind	1.78e-01		
22	6	30.0	265	2	S29119	chlorophyll a/b-bind	1.78e-01		
23	6	30.0	284	2	S68216	phosphoprotein phosph	1.78e-01		

97	5	25.0	77	2	PC1017	somatotropin precursor	4.27e+02	170	5	25.0	165	2	H70635	probable thiol peroxi	4.27e+02
98	5	25.0	79	2	A61326	lipotropin beta - ost	4.27e+02	171	5	25.0	166	1	S07369	interleukin-3 precurs	4.27e+02
99	5	25.0	80	2	T03404	metallothionein-like	4.27e+02	172	5	25.0	168	2	JC6566	interleukin-3 beta pr	4.27e+02
100	5	25.0	80	2	T03063	metallothionein-like	4.27e+02	173	5	25.0	169	2	T03033	probable cytochrome c	4.27e+02
101	5	25.0	82	2	T73096	MHC class II beta cha	4.27e+02	174	5	25.0	172	2	C71128	hypothetical protein	4.27e+02
102	5	25.0	83	2	T73095	MHC class II beta cha	4.27e+02	175	5	25.0	173	2	S09613	shikimate kinase (EC	4.27e+02
103	5	25.0	83	2	T02136	F8K4.11 - Arabidopsi	4.27e+02	176	5	25.0	176	2	A70516	hypothetical protein	4.27e+02
104	5	25.0	87	2	H50079	hypothetical protein	4.27e+02	177	5	25.0	178	2	E69389	hypothetical protein	4.27e+02
105	5	25.0	88	2	S70674	GUF 7-2a protein - co	4.27e+02	178	5	25.0	178	2	E65033	hypothetical 20.0 kD	4.27e+02
106	5	25.0	95	2	I53297	gene OR protein - rat	4.27e+02	179	5	25.0	179	2	S55660	capsid protein 65 - e	4.27e+02
107	5	25.0	95	2	S27146	gene 30.6 protein - p	4.27e+02	180	5	25.0	180	1	DXBHZ	protein 2 - barley (f	4.27e+02
108	5	25.0	99	2	D48362	hypothetical protein	4.27e+02	181	5	25.0	183	2	S22904	virion protein j13L -	4.27e+02
109	5	25.0	100	2	S15148	hypothetical protein	4.27e+02	182	5	25.0	185	2	S52903	virion protein j13L -	4.27e+02
110	5	25.0	101	2	S39201	1-aminocyclopropane-1	4.27e+02	183	5	25.0	186	1	WNZBA	matrix glycoprotein M	4.27e+02
111	5	25.0	104	2	S22706	vitelline membrane pr	4.27e+02	184	5	25.0	189	1	S52902	virion protein j13L -	4.27e+02
112	5	25.0	105	2	F69872	hypothetical protein	4.27e+02	185	5	25.0	189	1	WADM1	early E1A 20K protein	4.27e+02
113	5	25.0	106	1	R5BR1B	acidic ribosomal prot	4.27e+02	186	5	25.0	190	1	WMBE28	structural phosphopro	4.27e+02
114	5	25.0	106	2	S73034	hypothetical protein	4.27e+02	187	5	25.0	193	2	G69757	hypothetical protein	4.27e+02
115	5	25.0	109	1	R67E1T	acidic ribosomal prot	4.27e+02	188	5	25.0	194	2	S76452	hypothetical protein	4.27e+02
116	5	25.0	111	2	A64917	probable membrane pro	4.27e+02	189	5	25.0	194	2	A36037	yellow fluorescent pr	4.27e+02
117	5	25.0	111	2	H64159	hypothetical protein	4.27e+02	190	5	25.0	197	1	K1PGGU	guanylate kinase (EC	4.27e+02
118	5	25.0	111	2	C71401	hypothetical protein	4.27e+02	191	5	25.0	198	2	S39447	guanylate kinase (EC	4.27e+02
119	5	25.0	112	2	B47033	AadA2 - Pseudomonas a	4.27e+02	192	5	25.0	199	2	B70669	hypothetical protein	4.27e+02
120	5	25.0	112	2	S68584	hydrophobin A - culti	4.27e+02	193	5	25.0	200	2	S20445	GRP-binding protein,	4.27e+02
121	5	25.0	113	2	A60887	synuclein - Pacific e	4.27e+02	194	5	25.0	200	2	S49225	guanine nucleotide re	4.27e+02
122	5	25.0	118	2	B48099	transcription factor	4.27e+02	195	5	25.0	200	2	S23524	Grp-binding protein N	4.27e+02
123	5	25.0	119	2	S77863	ribosomal protein L17	4.27e+02	196	5	25.0	200	2	S23727	Grp-binding protein R	4.27e+02
124	5	25.0	119	2	E70161	ribosomal protein L18	4.27e+02	197	5	25.0	201	2	S66147	Grp-binding protein r	4.27e+02
125	5	25.0	120	2	S05791	matng pheromone alph	4.27e+02	198	5	25.0	203	2	S47691	hypothetical protein	4.27e+02
126	5	25.0	121	1	RIAFSI	somatostatin I precu	4.27e+02	199	5	25.0	204	2	F36950	urease accessory prot	4.27e+02
127	5	25.0	121	2	A27888	Ig heavy chain V regi	4.27e+02	200	5	25.0	205	2	I46876	phosphoprotein phosph	4.27e+02
128	5	25.0	122	1	FLHUTE	Ig heavy chain V-III	4.27e+02	201	5	25.0	205	2	S51301	hypothetical protein	4.27e+02
129	5	25.0	125	1	G1SP4	photosystem I chain, I	4.27e+02	202	5	25.0	206	2	S18250	collagen alpha 1(I)I	4.27e+02
130	5	25.0	126	2	A55539	hypothetical protein	4.27e+02	203	5	25.0	207	2	A41581	peptidylprolyl isomer	4.27e+02
131	5	25.0	129	2	S73260	ribosomal protein L12	4.27e+02	204	5	25.0	209	2	A36344	cortical cytoskeleton	4.27e+02
132	5	25.0	129	2	S37670	trab protein - Escher	4.27e+02	205	5	25.0	210	2	JC5512	SNARE protein 23 - mo	4.27e+02
133	5	25.0	130	2	S72926	hypothetical protein	4.27e+02	206	5	25.0	211	2	JS5296	vesicle-membrane fusi	4.27e+02
134	5	25.0	132	2	JQ0737	RnPA protein - Microc	4.27e+02	207	5	25.0	212	2	A64045	2-dehydro-3-deoxyphe	4.27e+02
135	5	25.0	132	2	I51211	drebrin E2 - chicken	4.27e+02	208	5	25.0	215	2	A59787	hypothetical protein	4.27e+02
136	5	25.0	135	3	JC2211	hypothetical 14.8K pr	4.27e+02	209	5	25.0	216	2	G70437	soluble hydrogenase s	4.27e+02
137	5	25.0	136	2	G70916	hypothetical protein	4.27e+02	210	5	25.0	216	2	S25172	pectinesterase (EC 3	4.27e+02
138	5	25.0	136	2	G70750	hypothetical protein	4.27e+02	211	5	25.0	218	2	H71015	hypothetical protein	4.27e+02
139	5	25.0	137	1	R5BY17	ribosomal protein L23	4.27e+02	212	5	25.0	220	2	S01838	nifv protein - Klebsi	4.27e+02
140	5	25.0	138	2	A69219	conserved hypothetica	4.27e+02	213	5	25.0	220	2	S48781	L1 protein - human pa	4.27e+02
141	5	25.0	138	2	T03374	high mobility group p	4.27e+02	214	5	25.0	223	2	S06705	hypothetical HTLV-1 r	4.27e+02
142	5	25.0	139	2	T03372	high mobility group p	4.27e+02	215	5	25.0	225	2	G70721	hypothetical protein	4.27e+02
143	5	25.0	140	1	R5RT23	ribosomal protein L23	4.27e+02	216	5	25.0	226	2	S41032	hypothetical protein	4.27e+02
144	5	25.0	140	2	S11937	histone H2B - Emerice	4.27e+02	217	5	25.0	227	2	S76546	hypothetical protein	4.27e+02
145	5	25.0	140	2	JC1253	ribosomal protein L17	4.27e+02	218	5	25.0	228	2	S65426	pyruvate decarboxylas	4.27e+02
146	5	25.0	141	2	F70457	hypothetical protein	4.27e+02	219	5	25.0	228	2	S55881	CCHH finger protein 1	4.27e+02
147	5	25.0	141	2	E32998	galactose-6-phosphate	4.27e+02	220	5	25.0	229	2	S24813	hypothetical protein	4.27e+02
148	5	25.0	142	2	A38158	mercury transport pro	4.27e+02	221	5	25.0	230	2	H64634	conserved hypothetica	4.27e+02
149	5	25.0	143	2	JS0479	probable general ser	4.27e+02	222	5	25.0	231	2	S34448	small nuclear ribonuc	4.27e+02
150	5	25.0	145	2	B61650	hypothetical protein	4.27e+02	223	5	25.0	231	2	C64826	aquaporin 2 - Escher	4.27e+02
151	5	25.0	145	2	S48428	hypothetical protein	4.27e+02	224	5	25.0	235	2	S55883	CCHH finger protein 3	4.27e+02
152	5	25.0	149	2	S50968	probable membrane pro	4.27e+02	225	5	25.0	236	1	R3BY10	ribosomal protein S6	4.27e+02
153	5	25.0	149	1	HSWT28	histone H2B.2 - wheat	4.27e+02	226	5	25.0	236	2	S49466	hypothetical protein	4.27e+02
154	5	25.0	151	2	S28048	histone H2B - maize	4.27e+02	227	5	25.0	237	2	PC4015	capsid protein VI pre	4.27e+02
155	5	25.0	152	2	S38249	hypothetical protein	4.27e+02	228	5	25.0	237	2	S37198	protein VI - mouse ad	4.27e+02
156	5	25.0	152	2	S22323	histone H2B - wheat	4.27e+02	229	5	25.0	237	2	A41044	octopine permease pro	4.27e+02
157	5	25.0	153	2	I73635	neuronal olfactomedin	4.27e+02	230	5	25.0	237	2	S40244	rab23 protein - mouse	4.27e+02
158	5	25.0	154	2	H70681	hypothetical protein	4.27e+02	231	5	25.0	238	2	H70866	hypothetical protein	4.27e+02
159	5	25.0	154	2	T02077	histone H2B - maize	4.27e+02	232	5	25.0	238	1	TWBYA1	protein N-acetyltrans	4.27e+02
160	5	25.0	155	1	CCRC50	cytochrome c2 precurs	4.27e+02	233	5	25.0	238	2	T01877	hypothetical protein	4.27e+02
161	5	25.0	155	2	G71184	probable 50S ribosoma	4.27e+02	234	5	25.0	238	2	E71406	probable carnitine ra	4.27e+02
162	5	25.0	155	2	S47797	hypothetical 17.5K pr	4.27e+02	235	5	25.0	240	2	A36791	hypothetical protein	4.27e+02
163	5	25.0	157	2	B69476	inosine monophosphate	4.27e+02	236	5	25.0	242	2	D64639	conserved hypothetica	4.27e+02
164	5	25.0	163	2	B69817	hypothetical protein	4.27e+02	237	5	25.0	242	2	T03538	uroporphyrin-III c-me	4.27e+02
165	5	25.0	160	2	I44020	TriB - plasmid RK2	4.27e+02	238	5	25.0	243	2	F70846	hypothetical protein	4.27e+02
166	5	25.0	161	2	T00741	acyl carrier protein,	4.27e+02	239	5	25.0	243	2	S16723	frcG protein - Methan	4.27e+02
167	5	25.0	161	2	A69732	PBSX prophage ORF xkd	4.27e+02	240	5	25.0	245	2	S03830	hypothetical protein	4.27e+02
168	5	25.0	163	2	F71547	hypothetical protein	4.27e+02	241	5	25.0	246	2	F70858	hypothetical protein	4.27e+02
169	5	25.0	163	1	UCLQ38	cuticle protein 38 -	4.27e+02	242	5	25.0	247	2	E64040	hypothetical protein	4.27e+02

243	5	25.0	251	2	A41768	orf 5' to uraA - Myxo	4.27e+02	316	5	25.0	303	2	G70756	hypothetical protein	4.27e+02
244	5	25.0	252	2	S3656	lipoteichoic acid bio	4.27e+02	317	5	25.0	303	2	A69542	conserved hypothetical	4.27e+02
245	5	25.0	253	2	F36914	molybdate-binding pro	4.27e+02	318	5	25.0	305	2	S70183	chemotaxis protein me	4.27e+02
246	5	25.0	254	2	F69181	transcription regulat	4.27e+02	319	5	25.0	305	2	JC5844	chitinase (EC 3.2.1.1	4.27e+02
247	5	25.0	253	2	H70780	hypothetical protein	4.27e+02	320	5	25.0	307	2	H70651	hypothetical protein	4.27e+02
248	5	25.0	254	2	F69309	ATP-binding protein P	4.27e+02	321	5	25.0	308	2	I48080	coatamer complex epsi	4.27e+02
249	5	25.0	255	2	A70758	hypothetical protein	4.27e+02	322	5	25.0	308	2	E71697	probable proteinase s	4.27e+02
250	5	25.0	256	2	E70868	probable enoyl-coA hy	4.27e+02	323	5	25.0	308	2	D70875	probable PE protein -	4.27e+02
251	5	25.0	256	1	TRFF	trypsin-like proteina	4.27e+02	324	5	25.0	308	2	A70761	hypothetical protein	4.27e+02
252	5	25.0	259	2	S70530	hypothetical protein	4.27e+02	325	5	25.0	309	2	G69726	tRNA pseudouridine 5S	4.27e+02
253	5	25.0	259	2	S55884	CCHH finger protein 4	4.27e+02	326	5	25.0	310	2	H71703	thioredoxin reductase	4.27e+02
254	5	25.0	260	2	A36299	transcription factor	4.27e+02	327	5	25.0	311	2	H71416	hypothetical protein	4.27e+02
255	5	25.0	261	2	S75631	cyclase hlsF - Synech	4.27e+02	328	5	25.0	313	2	A28444	flagellin precursor -	4.27e+02
256	5	25.0	261	2	S62047	probable membrane pro	4.27e+02	329	5	25.0	313	2	S46690	hypothetical protein	4.27e+02
257	5	25.0	261	2	S58496	IAAL protein - Arabid	4.27e+02	330	5	25.0	314	2	G69657	tRNA isopentenylprop	4.27e+02
258	5	25.0	261	2	S76510	hypothetical protein	4.27e+02	331	5	25.0	315	2	S32401	H+-transporting ATP s	4.27e+02
259	5	25.0	263	2	B42105	ribosomal protein MRP	4.27e+02	332	5	25.0	316	2	A71502	probable methionyl tr	4.27e+02
260	5	25.0	263	2	A35147	hypothetical protein	4.27e+02	333	5	25.0	317	2	I52575	gene MRK protein - mo	4.27e+02
261	5	25.0	263	2	T02227	NBS-LRR type resistan	4.27e+02	334	5	25.0	317	2	S48036	hypothetical protein	4.27e+02
262	5	25.0	263	2	B69906	rard protein homolog	4.27e+02	335	5	25.0	317	2	H69135	protein kinase - Meth	4.27e+02
263	5	25.0	264	2	D34768	ORF4 protein - Orf vi	4.27e+02	336	5	25.0	317	2	T00500	probable elicitor res	4.27e+02
264	5	25.0	266	2	A70718	hypothetical protein	4.27e+02	337	5	25.0	319	2	S07570	site-specific DNA-met	4.27e+02
265	5	25.0	267	2	F70967	hypothetical protein	4.27e+02	338	5	25.0	319	2	T01822	hypothetical protein	4.27e+02
266	5	25.0	268	2	S74440	iron(III) dicitrate t	4.27e+02	339	5	25.0	324	2	I81191	proteinase IV (EC 3.4	4.27e+02
267	5	25.0	269	2	B23957	streptomycin/spectino	4.27e+02	340	5	25.0	327	2	D64407	bifunctional short ch	4.27e+02
268	5	25.0	270	2	F70432	conserved hypothetical	4.27e+02	341	5	25.0	328	2	JS0459	gene H protein - phag	4.27e+02
269	5	25.0	270	2	T00557	probable trpA protein	4.27e+02	342	5	25.0	328	1	ZHBPFA	gene H protein - phag	4.27e+02
270	5	25.0	270	2	E69921	DNA ligase homolog yo	4.27e+02	343	5	25.0	328	2	S67592	hypothetical protein	4.27e+02
271	5	25.0	272	2	A71203	hypothetical protein	4.27e+02	344	5	25.0	329	2	T00873	hypothetical protein	4.27e+02
272	5	25.0	272	2	S52977	hypothetical protein	4.27e+02	345	5	25.0	330	2	H64894	tellurite resistance	4.27e+02
273	5	25.0	272	2	S55510	coat protein - potato	4.27e+02	346	5	25.0	331	2	B47236	zinc-finger protein P	4.27e+02
274	5	25.0	272	2	A43306	DnrI - Streptomycetes p	4.27e+02	347	5	25.0	332	2	S61567	hypothetical protein	4.27e+02
275	5	25.0	273	2	B23971	monophenol monooxygen	4.27e+02	348	5	25.0	333	2	G36788	hypothetical protein	4.27e+02
276	5	25.0	274	2	A64978	hypothetical protein	4.27e+02	349	5	25.0	333	2	S32114	85C protein - Mycobac	4.27e+02
277	5	25.0	276	2	S13585	nosy protein precursor	4.27e+02	350	5	25.0	333	2	A39065	homeotic protein EVX2	4.27e+02
278	5	25.0	276	2	F70340	acetyl-CoA carboxyltr	4.27e+02	351	5	25.0	333	2	A36925	transcription activat	4.27e+02
279	5	25.0	277	1	DC9615	orotidine-5'-phosphat	4.27e+02	352	5	25.0	334	2	F71114	probable ferredoxin o	4.27e+02
280	5	25.0	278	2	S29615	whiG protein - Strept	4.27e+02	353	5	25.0	334	2	C70813	hypothetical protein	4.27e+02
281	5	25.0	278	2	S27614	bromide peroxidase (E	4.27e+02	354	5	25.0	335	2	S65383	lysine--tRNA ligase (	4.27e+02
282	5	25.0	279	2	S52582	prephenate dehydratas	4.27e+02	355	5	25.0	335	2	E69990	proteinase IV homolog	4.27e+02
283	5	25.0	280	2	A33332	transcription initiat	4.27e+02	356	5	25.0	335	2	T00898	probable lytB' protei	4.27e+02
284	5	25.0	280	2	A61047	ectodermal (ect) - fr	4.27e+02	357	5	25.0	335	2	S42807	HSR203J protein - com	4.27e+02
285	5	25.0	281	1	ZBBE12	30.2k zinc-binding pr	4.27e+02	358	5	25.0	336	2	T00571	dolichyl-phosphate be	4.27e+02
286	5	25.0	281	2	JC4295	heat-shock protein -	4.27e+02	359	5	25.0	338	2	JL0121	glycerolaldehyde-3-phos	4.27e+02
287	5	25.0	281	2	T02813	hypothetical protein	4.27e+02	360	5	25.0	339	2	S20880	homeotic protein Hox	4.27e+02
288	5	25.0	282	2	A38536	Soj antagonist / chro	4.27e+02	361	5	25.0	339	2	D41377	moXR protein - Paraco	4.27e+02
289	5	25.0	282	2	D70207	antigen S2 - Lyme dis	4.27e+02	362	5	25.0	340	2	G64905	sugar-binding protein	4.27e+02
290	5	25.0	282	2	JT0613	catechol 1,2-dioxygen	4.27e+02	363	5	25.0	340	2	F70079	conserved hypothetical	4.27e+02
291	5	25.0	282	1	S59538	heat shock transcript	4.27e+02	364	5	25.0	341	2	H64854	probable glucosidase	4.27e+02
292	5	25.0	283	2	C70478	proteinase IV - Aquif	4.27e+02	365	5	25.0	342	2	A30189	iron stress-induced h	4.27e+02
293	5	25.0	283	2	S75226	esterase slr1916 - Sy	4.27e+02	366	5	25.0	343	2	B69800	iron(III) dicitrate t	4.27e+02
294	5	25.0	284	1	SYECOL	2-dehydro-3-deoxyphos	4.27e+02	367	5	25.0	345	2	G69138	conserved hypothetical	4.27e+02
295	5	25.0	284	2	C70018	opine catabolism homo	4.27e+02	368	5	25.0	345	2	D65023	lipoprotein-34 precu	4.27e+02
296	5	25.0	285	2	G71912	probable 3-hydroxyaci	4.27e+02	369	5	25.0	346	2	A58522	upstream stimulatory	4.27e+02
297	5	25.0	290	2	D70974	probable dehydrogenas	4.27e+02	370	5	25.0	346	2	A55111	transcription factor	4.27e+02
298	5	25.0	291	2	S37743	probable GTP-binding	4.27e+02	371	5	25.0	346	2	A34365	holocytochrome-c synt	4.27e+02
299	5	25.0	292	2	H70513	hypothetical protein	4.27e+02	372	5	25.0	347	2	S60652	synapse-associated pr	4.27e+02
300	5	25.0	293	2	A69251	N5-methyltetrahydra	4.27e+02	373	5	25.0	348	2	D24723	trpD protein - Coryne	4.27e+02
301	5	25.0	294	2	T00104	probable dUDP-4-dehyd	4.27e+02	374	5	25.0	349	2	A44507	licheninase (EC 3.2.1	4.27e+02
302	5	25.0	294	2	S60991	hypothetical protein	4.27e+02	375	5	25.0	349	2	C64875	probable serine prote	4.27e+02
303	5	25.0	296	2	S44972	lmbX protein - Strept	4.27e+02	376	5	25.0	349	2	T00526	probable proline-rich	4.27e+02
304	5	25.0	297	2	JL0032	hypothetical 31.7K pr	4.27e+02	377	5	25.0	351	2	H70570	hypothetical protein	4.27e+02
305	5	25.0	297	2	S55063	hypothetical protein	4.27e+02	378	5	25.0	351	2	JQ0148	hypothetical 34.4K pr	4.27e+02
306	5	25.0	298	2	S77536	lactose transport sys	4.27e+02	379	5	25.0	351	2	S60653	synapse associated pr	4.27e+02
307	5	25.0	298	2	A54860	p40 protein homolog -	4.27e+02	380	5	25.0	352	4	S25497	thymidine kinase (EC	4.27e+02
308	5	25.0	298	2	S71247	ribosome-associated p	4.27e+02	381	5	25.0	352	2	C37471	hypothetical helicase	4.27e+02
309	5	25.0	299	2	S56031	pathogenesis-related	4.27e+02	382	5	25.0	352	2	JS0637	RepB protein - Thioba	4.27e+02
310	5	25.0	299	2	S38896	hypothetical protein	4.27e+02	383	5	25.0	352	2	A36128	regulatory protein al	4.27e+02
311	5	25.0	299	2	PT0060	N-acetylphosphinothri	4.27e+02	384	5	25.0	352	1	KIBED2	thymidine kinase (EC	4.27e+02
312	5	25.0	299	2	A47031	bialaphos acetylhydro	4.27e+02	385	5	25.0	352	1	KIBEE4	thymidine kinase (EC	4.27e+02
313	5	25.0	300	2	I39495	hypothetical protein 4	4.27e+02	386	5	25.0	353	2	D64136	probable serine prote	4.27e+02
314	5	25.0	300	2	E71534	probable ppp2b methyl	4.27e+02	387	5	25.0	354	2	S61419	thiamine biosynthetic	4.27e+02
315	5	25.0	301	2	S35526	homeotic protein hoxd	4.27e+02	388	5	25.0	354	1	OWNS	ornithine carbamoyltr	4.27e+02

389	5	25.0	359	1	A24102	alanine racemase (EC	4.27e+02	462	5	25.0	406	2	A35360	otefin - fruit fly (D	4.27e+02
390	5	25.0	360	2	T02934	JUN-activation-domain	4.27e+02	463	5	25.0	407	2	I52703	42K membrane glycopro	4.27e+02
391	5	25.0	361	1	E54130	alanine racemase (EC	4.27e+02	464	5	25.0	408	2	A71657	dihydropolipamide acet	4.27e+02
392	5	25.0	361	2	S59548	1-aminocyclopropane-1	4.27e+02	465	5	25.0	408	2	G70925	probable PPE protein	4.27e+02
393	5	25.0	361	1	S31225	transcription factor	4.27e+02	466	5	25.0	408	2	JQ0814	yhad protein - Escher	4.27e+02
394	5	25.0	361	2	A48373	high-alkaline serine	4.27e+02	467	5	25.0	410	2	S70847	neuroserpin precursor	4.27e+02
395	5	25.0	361	2	S34613	5-exo-hydroxycamphor	4.27e+02	468	5	25.0	410	2	A64051	probable integral mem	4.27e+02
396	5	25.0	361	1	A55557	transcription factor	4.27e+02	469	5	25.0	410	2	S04391	translation elongatio	4.27e+02
397	5	25.0	362	1	E64793	glycerol dehydrogenas	4.27e+02	470	5	25.0	413	2	JC4763	capsid protein - Myco	4.27e+02
398	5	25.0	362	2	D70603	probable lipoprotein	4.27e+02	471	5	25.0	414	2	H64128	hypothetical protein	4.27e+02
399	5	25.0	363	2	I57955	sonatostatin receptor	4.27e+02	472	5	25.0	415	2	B70626	hypothetical protein	4.27e+02
400	5	25.0	364	2	JN0763	sonatostatin receptor	4.27e+02	473	5	25.0	415	2	S32932	regulatory protein sp	4.27e+02
401	5	25.0	364	1	D33675	thymidine kinase (EC	4.27e+02	474	5	25.0	415	2	A54350	probable transposase	4.27e+02
402	5	25.0	364	2	S28771	polylacturonase (EC	4.27e+02	475	5	25.0	415	2	H70351	probable IS1081 trans	4.27e+02
403	5	25.0	366	2	A46607	growth/differentiatio	4.27e+02	476	5	25.0	416	1	FOCH	transforming protein	4.27e+02
404	5	25.0	366	2	A45402	transforming growth f	4.27e+02	477	5	25.0	416	2	S52078	prostacyclin - rat	4.27e+02
405	5	25.0	367	1	DRECD4	aspartate-semialdehyd	4.27e+02	478	5	25.0	416	2	S27198	homeotic protein Hox	4.27e+02
406	5	25.0	367	2	I40837	2-dehydro-3-deoxyphos	4.27e+02	479	5	25.0	416	1	TVFVAC	transforming protein	4.27e+02
407	5	25.0	370	2	S32173	hypothetical protein	4.27e+02	480	5	25.0	416	2	C64985	hypothetical 43.4 kD	4.27e+02
408	5	25.0	371	1	A44144	transcription factor	4.27e+02	481	5	25.0	417	2	S57820	pyruvate decarboxylas	4.27e+02
409	5	25.0	372	2	C59417	citrate synthase (cit	4.27e+02	482	5	25.0	418	2	B65013	hypothetical protein	4.27e+02
410	5	25.0	372	1	I52196	homeobox transcriptio	4.27e+02	483	5	25.0	419	2	H70863	hypothetical protein	4.27e+02
411	5	25.0	372	2	A46037	Hox-1.11 - mouse	4.27e+02	484	5	25.0	419	2	S70817	homeotic protein goos	4.27e+02
412	5	25.0	373	2	S06010	finger protein egon -	4.27e+02	485	5	25.0	419	2	D64142	hypothetical protein	4.27e+02
413	5	25.0	373	1	S29934	chorismate mutase (EC	4.27e+02	486	5	25.0	422	2	S71243	3'-phosphoadenosine 5	4.27e+02
414	5	25.0	373	2	A59143	galactosyl-transferas	4.27e+02	487	5	25.0	423	2	G711315	probable DNA polymera	4.27e+02
415	5	25.0	373	1	KWECTD	chorismate mutase (EC	4.27e+02	488	5	25.0	423	2	JQ1975	chitinase (EC 3.2.1.1	4.27e+02
416	5	25.0	376	1	K1BET	thymidine kinase (EC	4.27e+02	489	5	25.0	423	1	TVFV2E	transforming protein	4.27e+02
417	5	25.0	376	2	JC6535	multubiquitin-chain-	4.27e+02	490	5	25.0	425	2	S41099	protein kinase (EC 2.	4.27e+02
418	5	25.0	376	1	K1BEB16	thymidine kinase (EC	4.27e+02	491	5	25.0	426	2	T03372	hypothetical protein	4.27e+02
419	5	25.0	376	1	K1BEB17	thymidine kinase (EC	4.27e+02	492	5	25.0	426	2	S63200	hypothetical protein	4.27e+02
420	5	25.0	376	1	K1BEBK5	thymidine kinase (EC	4.27e+02	493	5	25.0	427	1	A43698	paired box transcript	4.27e+02
421	5	25.0	376	2	S38661	aspartate-semialdehyd	4.27e+02	494	5	25.0	429	2	T01550	receptor kinase homol	4.27e+02
422	5	25.0	376	1	K1BEBHF	thymidine kinase (EC	4.27e+02	495	5	25.0	433	2	S51688	nucleocapsid protein	4.27e+02
423	5	25.0	376	1	K1BETC	thymidine kinase (EC	4.27e+02	496	5	25.0	434	2	B69271	hypothetical protein	4.27e+02
424	5	25.0	377	2	S33671	26S proteinase chain	4.27e+02	497	5	25.0	434	2	T02407	heme A farnesyltransf	4.27e+02
425	5	25.0	377	2	F69008	acetyltransferase - M	4.27e+02	498	5	25.0	438	2	T00507	indole-3-acetate beta	4.27e+02
426	5	25.0	377	2	B70874	probable moxr protein	4.27e+02	499	5	25.0	439	2	JC5253	TEA domain-containing	4.27e+02
427	5	25.0	378	2	JC4238	HMG-box transcription	4.27e+02	500	5	25.0	440	2	E711268	conserved hypothetica	4.27e+02
428	5	25.0	378	2	C64142	conserved hypothetica	4.27e+02	501	5	25.0	442	2	S57762	6-phospho-beta-glucos	4.27e+02
429	5	25.0	381	2	A64783	ybbz protein - Escher	4.27e+02	502	5	25.0	443	2	I38239	Sox3 - human	4.27e+02
430	5	25.0	382	1	A70071	conserved hypothetica	4.27e+02	503	5	25.0	443	2	PQ0450	tryptophan synthase (	4.27e+02
431	5	25.0	382	1	SYECLA	lipid-A-disaccharide	4.27e+02	504	5	25.0	444	2	C65102	hypothetical 49.0 kD	4.27e+02
432	5	25.0	383	2	A71139	hypothetical protein	4.27e+02	505	5	25.0	445	2	B57742	cyclin Ib - maize	4.27e+02
433	5	25.0	383	2	B48122	GTPase-activating pro	4.27e+02	506	5	25.0	445	2	S26601	transcription activat	4.27e+02
434	5	25.0	384	2	H64505	hypothetical protein	4.27e+02	507	5	25.0	446	2	S72743	4-aminobutyrate trans	4.27e+02
435	5	25.0	384	2	A48122	GTPase-activating pro	4.27e+02	508	5	25.0	446	2	G64227	glycine-tRNA ligase	4.27e+02
436	5	25.0	385	2	F70591	probable kefB protein	4.27e+02	509	5	25.0	446	2	S57900	conserved hypothetica	4.27e+02
437	5	25.0	387	2	S02222	probable chlorophyll	4.27e+02	510	5	25.0	446	2	S34997	pfs protein - pseudo	4.27e+02
438	5	25.0	388	2	G71142	hypothetical protein	4.27e+02	511	5	25.0	446	2	A41896	transcription activat	4.27e+02
439	5	25.0	389	2	A40809	enamelin, 44K - bovin	4.27e+02	512	5	25.0	447	2	B34714	retinoic acid recepto	4.27e+02
440	5	25.0	389	2	PQ0449	tryptophan synthase (	4.27e+02	513	5	25.0	448	2	A43786	retinoic acid recepto	4.27e+02
441	5	25.0	390	4	S03158	hypothetical BCR/ABL	4.27e+02	514	5	25.0	448	2	H70700	hypothetical protein	4.27e+02
442	5	25.0	390	2	C64996	hypothetical protein	4.27e+02	515	5	25.0	448	2	A57435	3,4-dihydroxyphenylac	4.27e+02
443	5	25.0	390	2	T02845	DDI1 protein - leishm	4.27e+02	516	5	25.0	448	2	S02827	retinoic acid recepto	4.27e+02
444	5	25.0	391	2	S75305	stage II sporulation	4.27e+02	517	5	25.0	448	2	S05051	retinoic acid recepto	4.27e+02
445	5	25.0	391	2	B70625	probable PPE protein	4.27e+02	518	5	25.0	451	1	TVFV2C	gag-myc polyprotein -	4.27e+02
446	5	25.0	391	2	S68354	NirF protein - Pseudo	4.27e+02	519	5	25.0	452	1	JNECGT	glycerol-3-phosphate	4.27e+02
447	5	25.0	394	2	S30286	tetracycline resistan	4.27e+02	520	5	25.0	454	2	E70503	hypothetical protein	4.27e+02
448	5	25.0	394	2	JC5197	aromatic-amino-acid t	4.27e+02	521	5	25.0	454	2	D70904	probable mceI protein	4.27e+02
449	5	25.0	395	2	JQ0430	hypothetical 44.4K pr	4.27e+02	522	5	25.0	454	2	S06124	retinoic acid recepto	4.27e+02
450	5	25.0	395	2	A54949	syndecan precursor -	4.27e+02	523	5	25.0	455	2	B70664	hypothetical protein	4.27e+02
451	5	25.0	396	1	S15992	flavohemoglobin hmp -	4.27e+02	524	5	25.0	455	2	S13512	retinoic acid recepto	4.27e+02
452	5	25.0	396	2	F70970	probable oxidoreducta	4.27e+02	525	5	25.0	455	2	PQ0407	coat protein - feline	4.27e+02
453	5	25.0	398	1	W2WL42	E2 protein - human pa	4.27e+02	526	5	25.0	456	2	D70772	hypothetical protein	4.27e+02
454	5	25.0	399	2	S13822	protein 24 - barley	4.27e+02	527	5	25.0	456	2	D71537	probable replication	4.27e+02
455	5	25.0	399	2	E54180	lipid-A-disaccharide	4.27e+02	528	5	25.0	458	2	A34714	retinoic acid recepto	4.27e+02
456	5	25.0	403	2	S76603	hypothetical protein	4.27e+02	529	5	25.0	458	2	C71420	hypothetical protein	4.27e+02
457	5	25.0	403	1	GRECY	tyrosine-specific tra	4.27e+02	530	5	25.0	459	2	S64924	XBj1 protein - yeast	4.27e+02
458	5	25.0	403	2	A53662	homeotic protein HB9	4.27e+02	531	5	25.0	459	2	E71443	probable DNA-binding	4.27e+02
459	5	25.0	403	2	H64463	hypothetical protein	4.27e+02	532	5	25.0	460	2	F70975	probable glutamate de	4.27e+02
460	5	25.0	405	2	S65471	pyruvate decarboxylas	4.27e+02	533	5	25.0	461	2	T03414	probable 1-aminocyclo	4.27e+02
461	5	25.0	405	2	C42094	bw3 protein - smut fu	4.27e+02	534	5	25.0	462	2	S30104	indoleacetamide hydro	4.27e+02



535	462	2	G70544	probable hemL protein	4.27e+02	5	25.0	508	1	WXHU	DNA nucleotidylexotra	4.27e+02
536	463	2	G69829	hypothetical protein	4.27e+02	5	25.0	509	1	DEFLP	dihydrolipoamide dehy	4.27e+02
537	463	2	D70933	hypothetical protein	4.27e+02	5	25.0	509	1	DEHULP	dihydrolipoamide dehy	4.27e+02
538	463	2	B70893	hypothetical glycine-	4.27e+02	5	25.0	509	2	H70597	probable membrane pro	4.27e+02
539	464	1	NDECKS	type I site-specific	4.27e+02	5	25.0	510	2	S38657	sucrose transport pro	4.27e+02
540	465	2	A43459	sperm membrane protei	4.27e+02	5	25.0	510	2	A55207	glycerol-3-phosphate	4.27e+02
541	466	2	I39707	hydrolase - Agrobacte	4.27e+02	5	25.0	512	2	D70506	hypothetical protein	4.27e+02
542	466	1	B43332	glutamate decarboxyla	4.27e+02	5	25.0	513	1	RGECAY	transcription regulat	4.27e+02
543	467	1	HJAGI	indoleacetamide hydro	4.27e+02	5	25.0	513	1	BMH06	bone morphogenetic pr	4.27e+02
544	467	2	A36738	DNA-binding protein T	4.27e+02	5	25.0	516	2	A24613	nicotinic acetylcholi	4.27e+02
545	467	2	S75150	glutamate decarboxyla	4.27e+02	5	25.0	516	2	S72937	hypothetical protein	4.27e+02
546	467	1	O2ACAT	indoleacetamide hydro	4.27e+02	5	25.0	517	2	A60916	nicotinic acetylcholi	4.27e+02
547	468	2	D70846	probable acyl-CoA deh	4.27e+02	5	25.0	517	2	S13875	nicotinic acetylcholi	4.27e+02
548	469	2	T01579	heme A:farnesyltransf	4.27e+02	5	25.0	518	2	T01318	pectinesterase 2 prec	4.27e+02
549	470	2	A31393	tryptophan synthase (	4.27e+02	5	25.0	519	2	S27396	modifier-3 protein -	4.27e+02
550	470	2	C70672	hypothetical protein	4.27e+02	5	25.0	519	2	S77572	oligopeptide transpor	4.27e+02
551	472	2	JC5199	nuclear hormone recep	4.27e+02	5	25.0	520	1	ACMSD1	nicotinic acetylcholi	4.27e+02
552	472	2	D70372	NADH oxidase - Aquife	4.27e+02	5	25.0	525	2	B43058	RNA-binding protein N	4.27e+02
553	474	2	C64801	yleA protein - Escher	4.27e+02	5	25.0	525	2	S27612	ketooglutarate semiald	4.27e+02
554	474	2	S75464	hypothetical protein	4.27e+02	5	25.0	525	2	S41763	histidine--trNA ligas	4.27e+02
555	474	2	S75032	hypothetical protein	4.27e+02	5	25.0	527	1	QRSEUB	sfiub protein - Serrat	4.27e+02
556	474	2	JX0301	ubiquinol--cytochrome	4.27e+02	5	25.0	528	2	T02863	hypothetical protein	4.27e+02
557	475	2	A43915	homeotic protein even	4.27e+02	5	25.0	529	2	B23595	DNA nucleotidylexotra	4.27e+02
558	475	2	JQ1073	tryptophan synthase (	4.27e+02	5	25.0	529	2	S62194	hypothetical protein	4.27e+02
559	477	1	ALBSN7	alpha-amylase (EC 3.2	4.27e+02	5	25.0	529	2	A65051	ygaa protein - Escher	4.27e+02
560	477	2	A47236	zinc-finger protein P	4.27e+02	5	25.0	531	2	A31340	glucuronosyltransfera	4.27e+02
561	477	1	WQEC2G	phosphotransferase sy	4.27e+02	5	25.0	531	2	T01415	basic leucine zipper	4.27e+02
562	477	1	TVMVCS	protein-tyrosine kina	4.27e+02	5	25.0	532	2	S65087	amino acid transporte	4.27e+02
563	477	2	A69144	succinoglycan biosynt	4.27e+02	5	25.0	534	2	S76219	hypothetical protein	4.27e+02
564	477	2	S36620	phosphotransferase sy	4.27e+02	5	25.0	535	2	F70914	probable long chain 1	4.27e+02
565	477	1	A36709	alpha-amylase (EC 3.2	4.27e+02	5	25.0	537	2	H71471	probable acylglycerol	4.27e+02
566	478	2	JC4838	bone morphogenetic pr	4.27e+02	5	25.0	540	3	FG0174	phospholipase D - Str	4.27e+02
567	478	2	JC1475	fatty-acyl-CoA reduct	4.27e+02	5	25.0	542	2	F64424	chaperonin - Methanoc	4.27e+02
568	478	1	S60754	transcription factor	4.27e+02	5	25.0	543	2	JQ1270	nucleoprotein - simia	4.27e+02
569	480	2	JN0558	UMP synthase - bovine	4.27e+02	5	25.0	544	2	S69072	conserved hypothetica	4.27e+02
570	480	2	T01207	cysteine proteinase m	4.27e+02	5	25.0	544	2	H70603	probable polyketide s	4.27e+02
571	481	2	H70679	probable membrane pro	4.27e+02	5	25.0	548	2	S52735	CW17R protein - mouse	4.27e+02
572	481	2	T01850	UTP-glucose glucosylt	4.27e+02	5	25.0	548	2	I68667	transcription factor	4.27e+02
573	482	2	G71345	conserved hypothetica	4.27e+02	5	25.0	548	2	A70780	probable membrane pro	4.27e+02
574	483	2	A31521	DNA-binding protein T	4.27e+02	5	25.0	548	2	S37969	probable triacylglyce	4.27e+02
575	484	2	S31721	xps2A protein - Xanth	4.27e+02	5	25.0	550	2	G70597	probable proteinase -	4.27e+02
576	485	1	RRYC62	RNA-directed DNA poly	4.27e+02	5	25.0	551	2	JC1111	H+-transporting ATP s	4.27e+02
577	485	2	F70464	hypothetical protein	4.27e+02	5	25.0	553	2	B55514	dihydrolipoamide S-ac	4.27e+02
578	485	2	I73637	neuronal olfactomedin	4.27e+02	5	25.0	553	2	E27793	methyl coenzyme M red	4.27e+02
579	485	2	D70664	probable gnd protein	4.27e+02	5	25.0	557	2	S41627	probable anti-mulleri	4.27e+02
580	485	2	B37855	glucose-6-phosphate 1	4.27e+02	5	25.0	558	2	S68981	aminonucleoside antib	4.27e+02
581	486	2	A64909	probable fructuronate	4.27e+02	5	25.0	558	2	A53719	glutamate dehydrogena	4.27e+02
582	488	2	S47072	finger protein H2F10,	4.27e+02	5	25.0	558	2	C70556	probable narH protein	4.27e+02
583	488	2	A69344	hypothetical protein	4.27e+02	5	25.0	559	2	A53280	alpha 1-adrenergic re	4.27e+02
584	490	2	S75539	hypothetical protein	4.27e+02	5	25.0	559	2	A38604	poly(3-hydroxyalkanoa	4.27e+02
585	490	2	H11480	probable GTPase/GTP-b	4.27e+02	5	25.0	559	2	S29305	poly(3-hydroxyalkanol	4.27e+02
586	490	2	S31450	1-aminocyclopropane-1	4.27e+02	5	25.0	560	2	A60164	platelet membrane gly	4.27e+02
587	490	2	A45623	Ankrd11.17 variant s	4.27e+02	5	25.0	560	2	C38604	poly(3-hydroxyalkanoa	4.27e+02
588	491	2	S58153	hypothetical protein	4.27e+02	5	25.0	560	2	A38731	alpha-1A-adrenergic r	4.27e+02
589	493	2	A41141	1-aminocyclopropane-1	4.27e+02	5	25.0	560	2	S29307	poly(3-hydroxyalkanol	4.27e+02
590	493	2	S71634	stearyl-CoA desaturat	4.27e+02	5	25.0	561	2	S17821	protoporphyrin IX mag	4.27e+02
591	493	2	A38649	1-aminocyclopropane-1	4.27e+02	5	25.0	561	2	A55235	rhodopy protein 2 pre	4.27e+02
592	494	2	A42170	zinc finger protein M	4.27e+02	5	25.0	564	2	B43776	drebrin E1 - chicken	4.27e+02
593	494	2	B41141	1-aminocyclopropane-1	4.27e+02	5	25.0	564	2	H70804	hypothetical protein	4.27e+02
594	494	2	E64077	probable glucose-6-ph	4.27e+02	5	25.0	564	2	B65011	hypothetical protein	4.27e+02
595	496	2	T01962	glutamate decarboxyla	4.27e+02	5	25.0	565	2	S04782	alpha,alpha-trehalase	4.27e+02
596	497	2	I70147	gene PP2A B protein -	4.27e+02	5	25.0	567	2	B69166	hypothetical protein	4.27e+02
597	497	2	JC5076	myc-associated zinc-f	4.27e+02	5	25.0	568	1	R32R1	ribosomal protein S1	4.27e+02
598	497	2	JC2192	subtilisin-like propr	4.27e+02	5	25.0	568	2	JC5629	muellerian-inhibiting	4.27e+02
599	498	2	S20518	probable serine/threo	4.27e+02	5	25.0	571	2	C70762	probable fatty-acid c	4.27e+02
600	499	2	I51257	retinoic acid recepto	4.27e+02	5	25.0	575	2	S35327	protein kinase seg39	4.27e+02
601	500	2	A48767	glutamate decarboxyla	4.27e+02	5	25.0	576	2	B70558	probable ABC transpor	4.27e+02
602	502	2	S61677	probable glutamate de	4.27e+02	5	25.0	579	2	S22584	rad4 protein - fission	4.27e+02
603	502	2	I52637	Ca2+/calmodulin-depen	4.27e+02	5	25.0	581	2	S53105	pectinesterase precu	4.27e+02
604	502	2	S41544	vsrA protein - Pseudo	4.27e+02	5	25.0	589	2	B29514	muscarinic acetylcholi	4.27e+02
605	506	2	H71253	probable glu-tRNA ami	4.27e+02	5	25.0	589	2	A23476	muscarinic acetylcholi	4.27e+02
606	507	2	S64507	probable membrane pro	4.27e+02	5	25.0	589	2	A34341	poly(3-hydroxybutyrat	4.27e+02
607	507	2	S48789	sucrose transport pro	4.27e+02	5	25.0	590	2	S01114	muscarinic acetylcholi	4.27e+02

681	5	25.0	590	2	S47572	muscarinic acetylcholin	4.27e+02	754	5	25.0	656	2	S76505	hypothetical protein	4.27e+02
682	5	25.0	590	2	S10128	muscarinic acetylcholin	4.27e+02	755	5	25.0	657	1	TVFFNA	transforming protein	4.27e+02
683	5	25.0	590	2	A44068	cell pattern formatio	4.27e+02	756	5	25.0	659	2	T02838	hypothetical protein	4.27e+02
684	5	25.0	590	1	WMCECB	64K capsid assembly p	4.27e+02	757	5	25.0	660	1	ALBS	alpha-amylase (EC 3.2	4.27e+02
685	5	25.0	591	2	E70802	hypothetical protein	4.27e+02	758	5	25.0	662	2	A33481	interferon-induced vi	4.27e+02
686	5	25.0	591	2	D70562	hypothetical protein	4.27e+02	759	5	25.0	662	2	A33481	interferon-induced vi	4.27e+02
687	5	25.0	592	2	JC4642	purH bifunctional enz	4.27e+02	760	5	25.0	667	2	Q02356	formate dehydrogenase	4.27e+02
688	5	25.0	592	1	ALBSNA	alpha-amylase (EC 3.2	4.27e+02	761	5	25.0	667	2	Q02356	formate dehydrogenase	4.27e+02
689	5	25.0	593	2	S55189	hypothetical protein	4.27e+02	762	5	25.0	675	2	G171640	capsid protein - fell	4.27e+02
690	5	25.0	594	2	I56248	receptor tyrosine kin	4.27e+02	763	5	25.0	675	2	S46952	NADH dehydrogenase I	4.27e+02
691	5	25.0	594	2	I58386	receptor tyrosine kin	4.27e+02	764	5	25.0	676	1	EDBE22	phosphotransferase sy	4.27e+02
692	5	25.0	595	2	B63893	propionyl-CoA carboxy	4.27e+02	765	5	25.0	676	1	EDBE22	phosphotransferase sy	4.27e+02
693	5	25.0	595	1	YFYBC	phenylalanine--tRNA l	4.27e+02	766	5	25.0	678	2	S54176	low affinity penicilli	4.27e+02
694	5	25.0	596	2	A55976	cellulose 1,4-beta-ce	4.27e+02	767	5	25.0	679	2	C71007	probable formate dehy	4.27e+02
695	5	25.0	596	2	I80177	TR4 orphan receptor -	4.27e+02	768	5	25.0	683	2	S69780	outer membrane protei	4.27e+02
696	5	25.0	596	2	JC4299	orphan nuclear recept	4.27e+02	769	5	25.0	686	2	S70180	chemotaxis protein ch	4.27e+02
697	5	25.0	597	2	S51212	BAK5 protein - bovine	4.27e+02	770	5	25.0	686	2	S70180	chemotaxis protein ch	4.27e+02
698	5	25.0	598	2	A57249	beta-galactosidase (E	4.27e+02	771	5	25.0	689	2	S26864	3-isopropylmalate deh	4.27e+02
699	5	25.0	600	2	F70980	probable accA3 protei	4.27e+02	772	5	25.0	693	1	TVHU42	bcr (breakpoint clust	4.27e+02
700	5	25.0	601	2	JC5720	vacuolar protein sort	4.27e+02	773	5	25.0	694	3	T00148	hypothetical protein	4.27e+02
701	5	25.0	602	2	S71557	pyruvate decarboxylas	4.27e+02	774	5	25.0	699	2	D70533	hypothetical protein	4.27e+02
702	5	25.0	603	2	B36738	DNA-binding protein T	4.27e+02	775	5	25.0	699	1	Q05374	US4 protein - human h	4.27e+02
703	5	25.0	605	2	A27274	ribophorin I precursor	4.27e+02	776	5	25.0	701	2	Q05374	US4 protein - human h	4.27e+02
704	5	25.0	605	2	G70040	sulfite reductase hom	4.27e+02	777	5	25.0	702	2	Q05374	US4 protein - human h	4.27e+02
705	5	25.0	607	2	A26168	ribophorin I precursor	4.27e+02	778	5	25.0	704	2	S51795	heat shock protein 90	4.27e+02
706	5	25.0	607	2	A43776	drebrin E2 - chicken	4.27e+02	779	5	25.0	704	2	T02902	hypothetical protein	4.27e+02
707	5	25.0	610	1	DC2MP	pyruvate decarboxylas	4.27e+02	780	5	25.0	706	2	D71289	probable single-stran	4.27e+02
708	5	25.0	610	1	S77177	proteinase IV (EC 3.4	4.27e+02	781	5	25.0	707	2	S60588	drebrin A - rat	4.27e+02
709	5	25.0	610	2	S77337	ABC-type transport pr	4.27e+02	782	5	25.0	707	1	TSBRAB	tryptophan synthase (	4.27e+02
710	5	25.0	610	2	S63667	phragmoplastin 12 - s	4.27e+02	783	5	25.0	708	2	B71558	probable low calcium	4.27e+02
711	5	25.0	610	2	S63668	phragmoplastin 5 - s	4.27e+02	784	5	25.0	710	2	H44490	retrovirus-related re	4.27e+02
712	5	25.0	611	1	S06047	endo-1,4-beta-xylanas	4.27e+02	785	5	25.0	713	2	S76766	hypothetical protein	4.27e+02
713	5	25.0	612	2	G64089	dihydroxy-acid dehydr	4.27e+02	786	5	25.0	715	2	S76492	lipoprotein nlpp - Sy	4.27e+02
714	5	25.0	613	2	A56031	potassium channel KCN	4.27e+02	787	5	25.0	716	2	JC5753	dishevelled protein 3	4.27e+02
715	5	25.0	614	2	B71551	probable s/t protein	4.27e+02	788	5	25.0	724	2	C49423	semaphorin II precurs	4.27e+02
716	5	25.0	615	2	F64128	proteinase IV homolog	4.27e+02	789	5	25.0	724	2	S29908	hypothetical protein	4.27e+02
717	5	25.0	615	1	S06977	nif-specific regulato	4.27e+02	790	5	25.0	724	1	WMVEPR	infected cell protein	4.27e+02
718	5	25.0	615	2	H64769	preprotein translocas	4.27e+02	791	5	25.0	725	1	WMVEPR	A-type inclusion prot	4.27e+02
719	5	25.0	615	2	I59309	TR4 orphan receptor -	4.27e+02	792	5	25.0	725	2	S13426	multidrug resistance	4.27e+02
720	5	25.0	616	1	DWECDA	dihydroxy-acid dehydr	4.27e+02	793	5	25.0	726	2	J02162	A-type inclusion prot	4.27e+02
721	5	25.0	618	1	PRCCTA	proteinase IV (EC 3.4	4.27e+02	794	5	25.0	728	2	S26427	structural protein, 7	4.27e+02
722	5	25.0	618	2	G69114	indolepyruvate oxidor	4.27e+02	795	5	25.0	728	2	A70632	hypothetical protein	4.27e+02
723	5	25.0	619	2	A23872	2-Isopropylmalate syn	4.27e+02	796	5	25.0	728	2	S71142	hypothetical protein	4.27e+02
724	5	25.0	619	2	B64064	heat shock cognate pr	4.27e+02	797	5	25.0	733	2	S44876	UC21.4 protein - Caen	4.27e+02
725	5	25.0	621	2	T01935	naringenin 3-dioxygen	4.27e+02	798	5	25.0	735	2	S46830	urea transport protei	4.27e+02
726	5	25.0	623	2	G70644	probable sppA protein	4.27e+02	799	5	25.0	738	2	S01892	hemolysin A precursor	4.27e+02
727	5	25.0	624	2	JC6023	poly(3-hydroxyalkanoa	4.27e+02	800	5	25.0	741	2	A41462	cytolysin/hemolysin p	4.27e+02
728	5	25.0	624	1	VCNW87	p87 capsid protein -	4.27e+02	801	5	25.0	742	2	A49341	isocitrate dehydrogen	4.27e+02
729	5	25.0	627	2	D69553	hypothetical protein	4.27e+02	802	5	25.0	751	2	T02858	hypothetical protein	4.27e+02
730	5	25.0	627	2	A44112	spidroin 2, dragline	4.27e+02	803	5	25.0	752	2	S61505	UDPglucose--starch gl	4.27e+02
731	5	25.0	627	2	T02846	hypothetical protein	4.27e+02	804	5	25.0	757	2	G71981	probable transcriptio	4.27e+02
732	5	25.0	627	2	T00124	hypothetical protein	4.27e+02	805	5	25.0	763	2	JC4376	beta-glucosidase (EC	4.27e+02
733	5	25.0	629	2	I54075	gene mR2R1 protein -	4.27e+02	806	5	25.0	767	2	I51083	SOX-L2 - rainbow trou	4.27e+02
734	5	25.0	631	2	JC2345	kinin-like proteinase	4.27e+02	807	5	25.0	769	2	H64525	transcription regulat	4.27e+02
735	5	25.0	633	2	D32053	parasporal crystal pr	4.27e+02	808	5	25.0	770	2	S11161	anthranilate synthase	4.27e+02
736	5	25.0	634	2	T00359	hypothetical protein	4.27e+02	809	5	25.0	770	2	B64892	ydbD protein - Escher	4.27e+02
737	5	25.0	634	2	A48335	cytoplasmic inclusion	4.27e+02	810	5	25.0	774	2	A39832	scabrous locus (scs)	4.27e+02
738	5	25.0	638	2	S22491	acetolactate synthase	4.27e+02	811	5	25.0	774	2	S58998	scabrous protein prec	4.27e+02
739	5	25.0	638	2	S22490	acetolactate synthase	4.27e+02	812	5	25.0	775	1	EDBE11	immediate-early prote	4.27e+02
740	5	25.0	639	2	T02784	calcium-dependent pro	4.27e+02	813	5	25.0	783	2	T00782	phosphoribosylanthran	4.27e+02
741	5	25.0	640	1	QCCHGM	phosphoenolpyruvate c	4.27e+02	814	5	25.0	784	2	A33633	transcription repress	4.27e+02
742	5	25.0	641	2	JC6017	zinc-finger transcrip	4.27e+02	815	5	25.0	785	2	S03785	dimethylsulfoxide red	4.27e+02
743	5	25.0	643	2	A41120	prostaglandin transpo	4.27e+02	816	5	25.0	789	1	A45617	6-phosphofructokinase	4.27e+02
744	5	25.0	643	2	S36017	finger protein tramsp	4.27e+02	817	5	25.0	789	2	S44759	C14B9.5 protein - Cae	4.27e+02
745	5	25.0	644	2	JC2346	kinin-like proteinase	4.27e+02	818	5	25.0	790	2	A35757	probable DNA-binding	4.27e+02
746	5	25.0	645	2	G01205	TYL protein - human	4.27e+02	819	5	25.0	799	2	A48716	glycine transporter G	4.27e+02
747	5	25.0	646	2	A70903	probable excludinase	4.27e+02	820	5	25.0	799	2	S65132	hypothetical protein	4.27e+02
748	5	25.0	647	2	PC1123	gag polyprotein - sim	4.27e+02	821	5	25.0	804	2	G54780	probable membrane pro	4.27e+02
749	5	25.0	651	2	S11735	hypothetical protein	4.27e+02	822	5	25.0	804	1	I5ECTB	DNA topoisomerase (AT	4.27e+02
750	5	25.0	652	2	S11735	resistance protein Mx	4.27e+02	823	5	25.0	806	2	A54078	DNA topoisomerase (AT	4.27e+02
751	5	25.0	652	2	JC2191	subtilisin-like propr	4.27e+02	824	5	25.0	806	2	H70647	probable NADH dehydro	4.27e+02
752	5	25.0	653	2	A32687	lysyl endopeptidase (	4.27e+02	825	5	25.0	808	2	A41538	Arp-binding cassette	4.27e+02
753	5	25.0	654	2	S53805	beta-glucosidase/xylo	4.27e+02	826	5	25.0	810	2	C70791	probable ponA' protei	4.27e+02

827	5	25.0	819	2	G69801	hypothetical protein	4.27e+02	900	1081	2	S51899	probable protein kina	4.27e+02
828	5	25.0	820	1	TVCIRF	protein-tyrosine kina	4.27e+02	901	1091	1	IJCHNL	neural cell adhesion	4.27e+02
829	5	25.0	821	2	F22930	DNA topoisomerase (AT	4.27e+02	902	1105	2	A71430	hypothetical protein	4.27e+02
830	5	25.0	822	1	TVHUFF	protein-tyrosine kina	4.27e+02	903	1133	1	EGRT	epidermal growth fact	4.27e+02
831	5	25.0	827	2	S59121	SOX6 protein - mouse	4.27e+02	904	1146	2	A55532	myosin-heavy-chain ki	4.27e+02
832	5	25.0	827	2	A36894	ATP-dependent protein	4.27e+02	905	1146	2	S46837	hypothetical protein	4.27e+02
833	5	25.0	827	2	A36895	endopeptidase Ia (EC	4.27e+02	906	1147	2	S47647	nitric-oxide synthase	4.27e+02
834	5	25.0	831	2	S26675	DNA-directed DNA poly	4.27e+02	907	1148	2	A28614	nonstructural polypro	4.27e+02
835	5	25.0	834	2	S74654	sensory transduction	4.27e+02	908	1168	1	MXAXIC	myosin heavy chain IC	4.27e+02
836	5	25.0	837	2	D69171	intracellular protein	4.27e+02	909	1170	2	A57650	repair protein XPG -	4.27e+02
837	5	25.0	841	2	S69563	suppressor protein PS	4.27e+02	910	1178	2	C71350	probable DNA-directed	4.27e+02
838	5	25.0	865	2	A47282	calcium-binding prote	4.27e+02	911	1181	2	A33998	integrin alpha-2 chal	4.27e+02
839	5	25.0	869	2	S72760	ppsl protein - Mycoba	4.27e+02	912	1185	2	S55552	LAR-interacting prote	4.27e+02
840	5	25.0	873	2	B53225	ecdysone-induced prot	4.27e+02	913	1194	2	E70940	probable COBN protein	4.27e+02
841	5	25.0	873	2	A47283	calphotin - fruit fly	4.27e+02	914	1194	2	D49851	magnesium-protoporph	4.27e+02
842	5	25.0	875	1	FOFVHB	gag-myc polyprotein -	4.27e+02	915	1196	2	S35994	DNA repair protein xp	4.27e+02
843	5	25.0	876	3	JE0283	HIRA protein - Fruit	4.27e+02	916	1201	2	S57369	anillin - fruit fly (	4.27e+02
844	5	25.0	887	1	W2BE95	gene 53 protein - equ	4.27e+02	917	1202	2	S55553	LAR-interacting prote	4.27e+02
845	5	25.0	894	2	F69730	cell wall-associated	4.27e+02	918	1206	2	S72620	probable reverse tran	4.27e+02
846	5	25.0	896	2	H69271	molybdopterin oxidore	4.27e+02	919	1216	2	A55620	apical endosomal prot	4.27e+02
847	5	25.0	898	2	T01503	hypothetical protein	4.27e+02	920	1248	2	A47445	reverse gyrase - Sulf	4.27e+02
848	5	25.0	901	2	G71386	probable pyruvate, ph	4.27e+02	921	1252	2	T00263	hypothetical protein	4.27e+02
849	5	25.0	905	2	A27410	plasma cell membrane	4.27e+02	922	1265	2	S57968	Ran-binding protein 2	4.27e+02
850	5	25.0	908	2	A33280	sarcalumenin precurs	4.27e+02	923	1271	1	TVHUBR	bcr (breakpoint clust	4.27e+02
851	5	25.0	912	2	F71433	probable growth regul	4.27e+02	924	1293	2	T01512	hypothetical protein	4.27e+02
852	5	25.0	918	1	I48719	protein kinase C (EC	4.27e+02	925	1298	1	EDBE75	immediate-early prote	4.27e+02
853	5	25.0	920	1	PXNCP	H+-transporting ATPas	4.27e+02	926	1309	2	S35484	peptidyl-diesteridase	4.27e+02
854	5	25.0	921	2	A48184	transcription initiat	4.27e+02	927	1317	2	B41950	retrovirus-related hy	4.27e+02
855	5	25.0	921	2	A45183	TBP-associated factor	4.27e+02	928	1319	2	D28313	glued protein - fruit	4.27e+02
856	5	25.0	923	2	A39596	progesterone receptor	4.27e+02	929	1327	2	D70759	probable otssb protein	4.27e+02
857	5	25.0	932	2	I52527	PACE4A - mouse (fragm	4.27e+02	930	1329	2	E70917	hypothetical glycine-	4.27e+02
858	5	25.0	936	2	S43738	transcription activat	4.27e+02	931	1364	2	S16845	suppressor protein -	4.27e+02
859	5	25.0	937	2	I53282	gene PACE4 protein -	4.27e+02	932	1365	2	S14871	senary transduction	4.27e+02
860	5	25.0	939	2	S18213	formate dehydrogenase	4.27e+02	933	1371	2	S77521	DNA-directed RNA poly	4.27e+02
861	5	25.0	948	2	A57640	retinoblastoma protei	4.27e+02	934	1386	2	G71529	hypothetical protein	4.27e+02
862	5	25.0	950	2	E64135	oxoglutarate dehydrog	4.27e+02	935	1408	2	S57049	hypothetical protein	4.27e+02
863	5	25.0	962	2	JC5571	subtilisin-like prote	4.27e+02	936	1421	2	T02501	hypothetical protein	4.27e+02
864	5	25.0	962	2	S03818	carboxymethylcellulas	4.27e+02	937	1438	2	B71610	WD40 WEB-1 homolog PF	4.27e+02
865	5	25.0	967	2	I40889	sarcosine oxidase (EC	4.27e+02	938	1442	2	S57160	sulfite reductase hom	4.27e+02
866	5	25.0	970	2	E70533	probable sulfatase -	4.27e+02	939	1464	2	S58984	development protein t	4.27e+02
867	5	25.0	972	2	A33926	DNA-directed RNA poly	4.27e+02	940	1487	1	EDBEF6	155K transcription ac	4.27e+02
868	5	25.0	975	2	JC5370	subtilisin-like prote	4.27e+02	941	1487	1	EDBEE1	immediate-early prote	4.27e+02
869	5	25.0	982	2	A44831	phosphoenolpyruvate c	4.27e+02	942	1489	2	S60416	DNA helicase YGL150c	4.27e+02
870	5	25.0	984	1	DJNVCP	probable DNA polymera	4.27e+02	943	1539	2	G70630	probable ctph protein	4.27e+02
871	5	25.0	985	1	VCLYSP	env polyprotein - hum	4.27e+02	944	1545	2	S71841	multidrug resistance	4.27e+02
872	5	25.0	985	1	VCLJSE	env polyprotein - sim	4.27e+02	945	1588	1	BVBVA1	AROL protein - yeast	4.27e+02
873	5	25.0	988	2	T03307	hypothetical protein	4.27e+02	946	1596	2	A35927	190K DNA-binding prot	4.27e+02
874	5	25.0	991	2	S43891	dhpl protein - fissio	4.27e+02	947	1603	1	BVASA1	arom protein - Emeric	4.27e+02
875	5	25.0	994	2	S18739	env protein - simian	4.27e+02	948	1607	2	T02837	long chain fatty acyl	4.27e+02
876	5	25.0	997	2	F71365	probable DNA polymera	4.27e+02	949	1615	3	JE0329	low density lipoprote	4.27e+02
877	5	25.0	997	2	S77625	mannuronan C-5-epime	4.27e+02	950	1615	3	JE0372	low density lipoprote	4.27e+02
878	5	25.0	1001	2	S74544	glycine dehydrogenase	4.27e+02	951	1646	1	WMTWS2	186K protein - cucumb	4.27e+02
879	5	25.0	1003	2	A39521	myosin IA - fruit fly	4.27e+02	952	1690	2	S41467	DNA-directed RNA poly	4.27e+02
880	5	25.0	1011	2	S45573	HEX2 protein - yeast	4.27e+02	953	1715	2	JE0128	Bombyx mori receptor	4.27e+02
881	5	25.0	1014	2	S32613	hypothetical protein	4.27e+02	954	1732	2	E71442	hypothetical protein	4.27e+02
882	5	25.0	1015	2	T00730	hypothetical protein	4.27e+02	955	1737	2	A37491	hypothetical helicase	4.27e+02
883	5	25.0	1025	1	A57676	protein kinase Xa21 (	4.27e+02	956	1749	2	S75138	hypothetical protein	4.27e+02
884	5	25.0	1030	2	I39887	lantibiotic subtilin	4.27e+02	957	1751	2	G71518	hypothetical protein	4.27e+02
885	5	25.0	1034	2	A24925	beta-galactosidase (E	4.27e+02	958	1752	2	S28849	DNA-directed RNA poly	4.27e+02
886	5	25.0	1034	2	JC5369	serine proteinase h2	4.27e+02	959	1785	2	A45546	major merizoite surfa	4.27e+02
887	5	25.0	1036	2	S55984	probable membrane pro	4.27e+02	960	1827	2	B70984	probable ppsd protein	4.27e+02
888	5	25.0	1036	2	JC5568	serine proteinase h1	4.27e+02	961	1839	2	S77626	mannuronan C-5-epime	4.27e+02
889	5	25.0	1044	2	S01966	GRPase-activating pro	4.27e+02	962	1849	2	T00415	hypothetical protein	4.27e+02
890	5	25.0	1048	1	XPBEA9	large structural phos	4.27e+02	963	1885	2	JC4086	fatty-acid synthase (	4.27e+02
891	5	25.0	1049	2	A34286	cytochrome P450 BM-3	4.27e+02	964	1956	2	T00051	hypothetical protein	4.27e+02
892	5	25.0	1049	1	IJHUG1	desmoglein 1 precurs	4.27e+02	965	2017	1	A36014	myosin II heavy chain, n	4.27e+02
893	5	25.0	1052	2	E71422	hypothetical protein	4.27e+02	966	2057	2	S61477	myosin II heavy chain	4.27e+02
894	5	25.0	1053	1	QYVC	phosphoenolpyruvate c	4.27e+02	967	2214	1	QZBYU2	pyrimidine synthesis	4.27e+02
895	5	25.0	1053	1	QYVC	phosphoenolpyruvate c	4.27e+02	968	2241	2	T02857	probable membrane pro	4.27e+02
896	5	25.0	1057	2	S72648	sucrose-phosphate syn	4.27e+02	969	2414	2	A54277	transcription adaptor	4.27e+02
897	5	25.0	1068	2	S64015	pleiotropic drug resi	4.27e+02	970	2431	1	MNVWSF	nonstructural polypro	4.27e+02
898	5	25.0	1071	2	S44798	F09G8.5 protein - Cae	4.27e+02	971	2479	1	MMVWRA	nonstructural polypro	4.27e+02
899	5	25.0	1077	2	T01474	hypothetical protein	4.27e+02	972	2555	2	C69681	peptide synthetase pp	4.27e+02

```
973 5 25.0 2560 2 I40457 peptide synthetase pp 4.27e+02
974 5 25.0 2594 2 A35774 kinase-related protei 4.27e+02
975 5 25.0 2671 2 A49873 inositol 1,4,5-tripho 4.27e+02
976 5 25.0 2693 2 A40743 IP3 receptor, XIP3R - 4.27e+02
977 5 25.0 2895 2 S54974 type 1 inositol 1,4,5 4.27e+02
978 5 25.0 2701 2 S17796 inositol-trisphosphat 4.27e+02
979 5 25.0 2713 2 A57113 inositol 1,4,5-tripho 4.27e+02
980 5 25.0 2723 2 T03221 probable polyketide s 4.27e+02
981 5 25.0 2734 2 B36579 inositol 1,4,5-tripho 4.27e+02
982 5 25.0 2749 2 A36579 inositol 1,4,5-tripho 4.27e+02
983 5 25.0 2749 1 ACMSIT inositol 1,4,5-tripho 4.27e+02
984 5 25.0 2833 2 A43360 inositol 1,4,5-trishp 4.27e+02
985 5 25.0 3005 1 GNVSTV genome polyprotein - 4.27e+02
986 5 25.0 3027 2 JQ1917 polyprotein - parsnip 4.27e+02
987 5 25.0 3063 2 JS0166 genome polyprotein - 4.27e+02
988 5 25.0 3224 2 S59884 Ran-binding protein 2 4.27e+02
989 5 25.0 3229 2 S27852 probable cell-surface 4.27e+02
990 5 25.0 3436 2 S36659 tegument protein 64 - 4.27e+02
991 5 25.0 3573 2 S23070 erythronolide synthas 4.27e+02
992 5 25.0 3649 2 A35085 trithorax protein - f 4.27e+02
993 5 25.0 3759 2 S18268 delta-(L-alpha-aminoa 4.27e+02
994 5 25.0 3770 2 A40889 delta-(L-alpha-aminoa 4.27e+02
995 5 25.0 3866 2 B48205 All-1 protein -GFE fo 4.27e+02
996 5 25.0 3869 2 A48205 All-1 protein +GFE fo 4.27e+02
997 5 25.0 3910 2 A44264 trithorax homolog HFX 4.27e+02
998 5 25.0 3968 2 A44265 trithorax homolog HFX 4.27e+02
999 5 25.0 4488 1 RRTHM2 genome polyprotein la 4.27e+02
1000 5 25.0 5126 2 S40450 ryanodine receptor/ca 4.27e+02
```

ALIGNMENTS

```
1
RESULT 1
ENTRY   S66279      #type complete
TITLE   proline dehydrogenase (EC 1.5.99.8) /
        1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12) -
        Salmonella typhimurium
CONTAINS
        1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12);
        proline dehydrogenase (EC 1.5.99.8)
ORGANISM #formal_name Salmonella typhimurium
DATE     12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
        17-Mar-1999
ACCESSIONS S66279; S33716; B39192; S03817; S31910
REFERENCE   S66279
#authors   Maloy, S.R.
#submission submitted to the EMBL Data Library, April 1994
#accession S66279
#molecule_type DNA
#residues 1-1320 #label MAL
##cross-references EMBL:X70843; NID:g470179; PID:g470180
##experimental_source strain LT2
#note      this is a revision to the sequence from reference S33716
REFERENCE   S33716
#authors   Allen, S.W.; Senti-Willis, A.; Maloy, S.R.
#journal   Nucleic Acids Res. (1993) 21:1676
#title     DNA sequence of the putA gene from Salmonella typhimurium: a
        bifunctional membrane-associated dehydrogenase that binds
        DNA.
#cross-references EMBL:X70843
#note         the nucleotide sequence was submitted to the EMBL Data
        Library, January 1993
#note         this sequence has been revised in reference S66279
REFERENCE   S53664
```

```
#authors   Ling, M.; Allen, S.W.; Wood, J.M.
#journal   J. Mol. Biol. (1994) 243:950-956
#title     Sequence analysis identifies the proline dehydrogenase and
        Delta(1)-pyrroline-5-carboxylate dehydrogenase domains of
        the multifunctional Escherichia coli PutA protein.
#cross-references MUID:95055736
#contents   annotation
REFERENCE   A39192
#authors   Ostrovsky de Spicer, P.; O'Brien, K.; Maloy, S.
#journal   J. Bacteriol. (1991) 173:211-219
#title     Regulation of proline utilization in Salmonella typhimurium:
        a membrane-associated dehydrogenase binds DNA in vitro.
#cross-references MUID:91100285
#accession B39192
#status     preliminary
#molecule_type DNA
#residues 1-26 #label OST
#note       the authors translated the codon CGT for residue 24 as
        Gly
REFERENCE   S03816
#authors   Hahn, D.R.; Myers, R.S.; Kent, C.R.; Maloy, S.R.
#journal   Mol. Gen. Genet. (1988) 213:125-133
#title     Regulation of proline utilization in Salmonella typhimurium:
        molecular characterization of the put operon, and DNA
        sequence of the put control region.
#cross-references MUID:89127131
#accession S03817
#molecule_type DNA
#residues 1-13,'HARTDQ',20-26 #label HAH
##cross-references EMBL:X12569
GENETICS
#gene       putA
#map_position 22 min
FUNCTION     #description transfers eletrons from proline to the respiratory chain;
        catalyzes proline oxidation to 1-pyrroline-5-carboxylate
        proline utilization
        membrane-bound with proline
        <HPC>
        #description catalyzes hydrolysis of 1-pyrroline-5-carboxylate to
        gamma-glutamic semialdehyde
        proline utilization
        <CPA>
        #description transcriptional repressor controls expression of genes putP
        and putA in response to proline supply
        #superfamily bifunctional protein putA
        DNA binding; FAD; flavoprotein; membrane-associated protein;
        NAD; oxidoreductase; transcription regulation
FEATURE      883      #active_site Glu #status predicted\
        917      #active_site Cys #status predicted\
SUMMARY      #length 1320 #molecular-weight 144187 #checksum 3619
Query Match 40.0%; Score 8; DB 2; Length 1320;
Best Local Similarity 100.0%; Pred. NO. 1.10e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1228 LAKRLPAA 1235
Qy 9 LAKRLPAA 16
|||||||
RESULT 2
ENTRY vesicular integral membrane protein VIP17 - dog
TITLE proteolipid VIP17; T-cell-specific MAL protein homolog VIP17
ALTERNATE_NAMES #formal_name Canis lupus familiaris common_name dog
ORGANISM #formal_name Canis lupus familiaris common_name dog
DATE 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change
        18-Sep-1998
ACCESSIONS S68406; S68399
REFERENCE S68399
#authors Zacchetti, D.; Peraenen, J.; Murata, M.; Fiedler, K.; Simons,
        K.
```

```

#journal      FEBS Lett. (1995) 377:465-469
#title        VIP17/MAL, a proteolipid in apical transport vesicles.
#accession    S68406
#status       nucleic acid sequence not shown
#molecule_type mRNA
#residues     1-153 ##label ZAC
##cross-references EMBL:X92508; NID:g1054712; PID:g1054713
#accession    S68399
#molecule_type protein
#residues     3-16;120-131 ##label ZAW
##experimental_source kidney; cell-line Madin-Darby canine kidney cell
FUNCTION
#pathway      myelin biogenesis
#superfamily   T-cell surface glycoprotein MAL
CLASSIFICATION #length 153 #molecular-weight 16624 #checksum 2815
SUMMARY
Query Match      35.0%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 5.09e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 PAAASGG 9
    |||||
QY 14 PAAASGG 20

RESULT 3
ENTRY      #type fragment
TITLE      transferrin binding protein 2 - Neisseria meningitidis
            (fragment)
ORGANISM   #formal_name Neisseria meningitidis
#variety   strain 528 #sequence_revision 13-Mar-1997 #text_change
            17-Mar-1999
ACCESSIONS S61547
REFERENCE   S61544
#authors    Rokbi, B.; Maitre-Wilmotte, G.; Mazarin, V.; Fourrichon, L.;
            Lissolo, L.; Quentin-Willet, M.J.;
            FEMS Microbiol. Lett. (1995) 132:277-283
#journal    Variable sequences in a mosaic-like domain of meningococcal
#title      tbp2 encode immunoreactive epitopes.
#cross-references MUID:96039602
#accession  S61547
#status     preliminary; translation not shown
##molecule_type DNA
##residues 1-35 ##label ROK
##cross-references EMBL:X88868
CLASSIFICATION #superfamily bacterial pathogen transferrin-binding protein;
            tonB-dependent receptor amino-terminal homology;
            tonB-dependent receptor carboxyl-terminal homology
KEYWORDS      membrane protein
SUMMARY       #length 35 #checksum 6464

Query Match      30.0%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.78e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 AAASGG 16
    |||||
QY 15 AAASGG 20

RESULT 4
ENTRY      #type complete
TITLE      acidic ribosomal protein p2.beta (X) - Trypanosoma cruzi
ALTERNATE_NAMES acidic ribosomal protein P.JF5 variant C
ORGANISM     #formal_name Trypanosoma cruzi
DATE         24-Aug-1996 #sequence_revision 08-Nov-1996 #text_change
            17-Mar-1999
ACCESSIONS S59919; S31433
REFERENCE   S59912
#authors    Schijman, A.G.; Vazquez, M.P.; Ben Dov, C.; Ghio, S.;
            Lorenza, H.; Levin, M.J.;
            Biochim. Biophys. Acta (1995) 1264:15-18

#journal      Cloning and sequence analysis of the TcP2-beta cDNA variants
            of Trypanosoma cruzi.
#accession    S59913
#status       not compared with conceptual translation
##molecule_type mRNA
##residues 1-107 ##label SCH
CLASSIFICATION #superfamily rat acidic ribosomal protein P1
KEYWORDS      phosphoprotein; protein biosynthesis; ribosome
SUMMARY       #length 107 #molecular-weight 10409 #checksum 2356

Query Match      30.0%; Score 6; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.78e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AAASGG 85
    |||||
QY 15 AAASGG 20

RESULT 6
ENTRY      #type complete
TITLE      acidic ribosomal protein P2 - Alternaria alternata
ORGANISM     #formal_name Alternaria alternata
DATE         19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change
            08-Sep-1997
ACCESSIONS S43109
REFERENCE   S43108
#authors    Achatz, G.; Oberkofler, H.; Simon, B.; Lechenauer, E.; Unger,
            A.; Kandler, D.; Prillinger, H.J.; Ebner, C.; Kraft, D.;
            Breitenbach, M.
#submission  submitted to the EMBL Data Library, March 1994
#description Molecular characterization of allergens of Cladosporium
            herbarum and Alternaria alternata.
#accession  S43109
##molecule_type mRNA
##residues 1-113 ##label ACH
##cross-references EMBL:X78222; NID:g467616; PID:g467617
CLASSIFICATION #superfamily rat acidic ribosomal protein P1
SUMMARY       #length 113 #molecular-weight 11134 #checksum 5473

Query Match      30.0%; Score 6; DB 2; Length 113;

```

```

#title        Cloning and sequence analysis of the TcP2-beta cDNA variants
            of Trypanosoma cruzi.
#cross-references MUID:96038812
#accession    S59919
##molecule_type mRNA
##residues 1-107 ##label SCH
##cross-references EMBL:X9509; NID:g10637; PID:g10638
CLASSIFICATION #superfamily rat acidic ribosomal protein P1
KEYWORDS      phosphoprotein; protein biosynthesis; ribosome
SUMMARY       #length 107 #molecular-weight 10421 #checksum 1314

Query Match      30.0%; Score 6; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.78e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AAASGG 85
    |||||
QY 15 AAASGG 20

RESULT 5
ENTRY      #type complete
TITLE      acidic ribosomal protein P2.beta (H1.3) - Trypanosoma cruzi
ORGANISM     #formal_name Trypanosoma cruzi
DATE         24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change
            17-Mar-1999
ACCESSIONS S59913
REFERENCE   S59912
#authors    Schijman, A.G.; Vazquez, M.P.; Ben Dov, C.; Ghio, S.;
            Lorenza, H.; Levin, M.J.;
            Biochim. Biophys. Acta (1995) 1264:15-18
#journal    Cloning and sequence analysis of the TcP2-beta cDNA variants
            of Trypanosoma cruzi.
#title      #cross-references MUID:96038812
#accession  S59913
#status       not compared with conceptual translation
##molecule_type mRNA
##residues 1-107 ##label SCH
CLASSIFICATION #superfamily rat acidic ribosomal protein P1
KEYWORDS      phosphoprotein; protein biosynthesis; ribosome
SUMMARY       #length 107 #molecular-weight 10409 #checksum 2356

Query Match      30.0%; Score 6; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.78e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AAASGG 85
    |||||
QY 15 AAASGG 20

```

```

Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 73 AAASGG 78
Qy 15 AAASGG 20

RESULT 7
ENTRY acidc ribosomal protein p3a - maize
TITLE #formal_name Zea mays #common_name maize
ORGANISM 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change
DATE 26-Feb-1999
ACCESSIONS T02037
REFERENCE 214507
#authors Bailey-Serres, J.; Vangala, S.; Szick, K.; Lee, C.H.
#journal Plant Physiol. (1997) 114:1293-1305
#title Acidic phosphoprotein complex of the 60S ribosomal subunit of
maize seedling roots. Components and changes in response to
flooding.
#accession T02037
#status preliminary; translated from GB/EMBL/DDBJ
#molecule_type mRNA
#residues 1-120 #label BAI
#cross-references EMBL:U62751; NID:g2431766; PID:g2431767
#experimental_source strain B73
GENETICS rpp3a
#gene protein biosynthesis; ribosome
KEYWORDS #length 120 #molecular-weight 12219 #checksum 6592
SUMMARY
Query Match 30.0%; Score 6; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 AAASGG 92
Qy 15 AAASGG 20

RESULT 8
ENTRY RHUN
TITLE #type complete
CONTAINS galanin precursor - human
CONTAINS galanin; galanin mRNA associated peptide (GMAP)
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Dec-1992 #sequence_revision 23-Mar-1995 #text_change
ACCESSIONS A49353; A61116; A49154; A41630; S15831
REFERENCE A49353
#authors Evans, H.; Baumgartner, M.; Shine, J.; Herzog, H.
#journal Genomics (1993) 18:473-477
#title Genomic organization and localization of the gene encoding
human preprogalanin.
#cross-references MUID:94140342
#accession A49353
#molecule_type DNA
#residues 1-45 #label EV2
#cross-references GB:L11144; NID:g306764; PID:g48849
#note neither nucleotide sequence nor translation is complete
REFERENCE A61116
#authors Evans, H.F.; Shine, J.
#journal Endocrinology (1991) 129:1682-1684
#title Human galanin: molecular cloning reveals a unique structure.
#cross-references MUID:91339793
#accession A61116
#molecule_type mRNA
#residues 1-123 #label EVA
REFERENCE A49154
#authors McKnight, G.L.; Karlsen, A.E.; Kowalyk, S.; Mathewes, S.L.;
Sheppard, P.O.; O'Hara, P.J.; Taborsky Jr., G.J.
#journal Diabetes (1992) 41:82-87
#title Sequence of human galanin and its inhibition of

```

```

glucose-stimulated insulin secretion from RIN cells.
#cross-references MUID:92090542
#accession A49154
#molecule_type mRNA
#residues 17-123 #label MCK
#cross-references GB:M77140; NID:g190411; PID:g190412
#note sequence extracted from NCBI backbone (NCBIN:71608,
NCBIP:71609)

REFERENCE A41630
#authors Schmidt, W.E.; Kratzin, H.; Eckart, K.; Drevs, D.;
Mundkowski, G.; Clemens, A.; Katsoulis, S.; Schaefer, H.;
Gallwitz, B.; Creutzfeldt, W.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:11435-11439
#title Isolation and primary structure of pituitary human galanin, a
30-residue nonamidated neuropeptide.
#cross-references MUID:92107965
#accession A41630
#molecule_type protein
#residues 33-62 #label SCH
#note sequence extracted from NCBI backbone
REFERENCE S15831
#authors Bersani, M.; Johnsen, A.H.; Hojrup, P.; Dunning, B.E.;
Andreasen, J.J.; Holst, J.J.
#journal FEBS Lett. (1991) 283:189-194
#title Human galanin: primary structure and identification of two
molecular forms.
#cross-references MUID:91257299
#accession S15831
#molecule_type protein
#residues 33-62 #label BER
#note a form containing only the first nineteen residues of
this sequence was also found
COMMENT This peptide causes smooth muscle contraction in the
gastrointestinal tract and genitourinary tract, inhibits insulin
release, and regulates growth hormone release.
COMMENT This peptide in human differs from the pig, rat, and bovine forms
by the absence of amidation at its carboxyl end and the presence
instead of an additional amino acid.
GENETICS GDB:GALN
#gene
#cross-references GDB:141567; OMIM:137035
#map_position 11q13.3-11q13.5
CLASSIFICATION #superfamily galanin
KEYWORDS hormone; intestine; neuropeptide
FEATURE
1-23 #domain signal sequence #status predicted #label SIG\
33-62 #product galanin #status experimental #label MAT\
65-123 #product galanin mRNA associated peptide (GMAP) #status
predicted #label RMAT
SUMMARY #length 123 #molecular-weight 13302 #checksum 678
Query Match 30.0%; Score 6; DB 1; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 111 LPAAS 116
Qy 13 LPAAS 18

RESULT 9
ENTRY #type complete
TITLE fimbrial protein precursor - Pseudomonas aeruginosa (strain
K122-4)
ORGANISM #formal_name Pseudomonas aeruginosa
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
ACCESSIONS A31105
REFERENCE A91879
#authors Pasloske, B.L.; Sastry, P.A.; Finlay, B.B.; Paranchych, W.
#journal J. Bacteriol. (1988) 170:3738-3741
#title Two unusual pilin sequences from different isolates of
Pseudomonas aeruginosa.

```

```

#cross-references MUID:88298689
#accession A31105
#molecule_type DNA
#residues 1-157 ##label PAS
##cross-references GB:M21651; NID:g151473; PID:g151474
CLASSIFICATION #superfamily gonococcal fimbrial protein
KEYWORDS methylated amino end
FEATURE
1-7 #domain propeptide #status predicted #label PRO\
8 #product fimbrial protein #status predicted #label MAT\
8 #modified_site methylated amino end (Phe) (in mature
form) #status predicted
SUMMARY #length 157 #molecular-weight 16175 #checksum 3862

Query Match 30.0%; Score 6; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 AAASGG 99
111111
QY 15 AAASGG 20

RESULT 10
ENTRY S22348 #type complete
TITLE H+-transporting ATP synthase (EC 3.6.1.34) delta chain
precursor - Human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
19-Feb-1999
ACCESSIONS S22348; T00493
REFERENCE Jordan, E.M.; Breen, G.A.M.
Biochim. Biophys. Acta (1992) 1130:123-126
Molecular cloning of an import precursor of the delta-subunit
of the human mitochondrial ATP synthase complex.
#cross-references MUID:92182007
#accession S22348
#status preliminary
#molecule_type mRNA
#residues 1-168 ##label JOR
##cross-references EMBL:X63422; NID:g12585; PID:g12586
REFERENCE Z14157
#authors Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson,
A.W.; Burkhart-Schultz, K.; Gordon, L.; Kyle, A.; Ramirez,
M.; Stillwagen, S.; Phan, H.; Velasco, N.; Garnes, J.;
Danganan, L.; Poundstone, P.; Christensen, M.; Georgescu,
A.; Avila, J.; Liu, S.; Attix, C.; Andreise, T.; Trankheim,
M.; Amico-Keller, G.; Coefield, J.; Duarte, S.; Lucas, S.;
Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Arellano,
A.; Montgomery, M.; Ow, D.; Nolan, M.; Trong, S.;
Kobayashi, A.; Olsen, A.O.; Carrano, A.V.
#submission submitted to the EMBL Data Library, February 1998
#description Sequence analysis of a 3.5 Mb contig in human 19p13.3
containing a serine protease gene cluster.
#accession T00493
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-168 ##label LAM
##cross-references EMBL:AC004221; NID:g2911257; PID:g2911258
GENETICS
#gene ATP5D
#map_position 19
#introns 47/3; 99/1; 128/3
the ATP synthase F1 complex consists of three alpha chains
(see PIR:PWHA), three beta chains (see PIR:A33370), and one
gamma chain (see PIR:A9108), one delta chain, and one
epsilon chain; the F1 complex binds to the intrinsic
membrane F0 complex
FUNCTION
#description catalyzes the formation of ATP from ADP and phosphate using
the free energy derived from proton transport down the
gradient maintained by cytochrome-c oxidase across the

```

```

#pathway mitochondrial inner-membrane
CLASSIFICATION oxidative phosphorylation
KEYWORDS #superfamily H+-transporting ATP synthase epsilon chain
hydroxylase; mitochondrion
SUMMARY #length 168 #molecular-weight 17490 #checksum 1686

Query Match 30.0%; Score 6; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 29 PAAASG 34
111111
QY 14 PAAASG 19

RESULT 11
ENTRY WMBEF5 #type complete
TITLE UL45 protein - human herpesvirus 1 (strain 17)
ORGANISM #formal_name human herpesvirus 1
DATE 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change
05-Sep-1997
ACCESSIONS I30088
REFERENCE McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.;
Frame, M.C.; McNab, D.; Perry, L.J.; Scott, J.E.; Taylor,
P.
J. Gen. Virol. (1988) 69:1531-1574
#journal The complete DNA sequence of the long unique region in the
genome of herpes simplex virus type 1.
#cross-references MUID:88274327
#accession I30088
#status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
#residues 1-172 ##label MCG
##cross-references GB:X14112; NID:gi944536; PID:g312360; PID:g59545;
GB:D00317
GENETICS
#gene UL45
CLASSIFICATION #superfamily herpesvirus UL45 protein
SUMMARY #length 172 #molecular-weight 18180 #checksum 8678

Query Match 30.0%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 RLPAAG 28
111111
QY 12 RLPAAG 17

RESULT 12
ENTRY S06613 #type complete
TITLE chorion protein sl9 - fruit fly (Drosophila grimshawi)
ORGANISM #formal_name Drosophila grimshawi
DATE 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
24-Sep-1998
ACCESSIONS S06613
REFERENCE S06612
#authors Martinez-Cruzado, J.C.; Swimmer, C.; Fenerjian, M.G.;
Kafatos, F.C.
#journal Genetics (1988) 119:663-677
#title Evolution of the autosomal chorion locus in Drosophila. I.
General organization of the locus and sequence comparisons
of genes sl5 and sl9 in evolutionarily distant species.
#cross-references MUID:88297142
#accession S06613
#molecule_type DNA
#residues 1-196 ##label MAR
##cross-references GB:X53422; EMBL:X12635; NID:g7419; PID:g7422
GENETICS
#gene sl9-1
#cross-references FlyBase:FBgn0012320
#map_position 5

```

```

#introns
SUMMARY
#length 196 #molecular-weight 19826 #checksum 4664
Query Match 30.0%; Score 6; DB 2; Length 196;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 AAASGG 61
Qy 15 AAASGG 20

RESULT 13
ENTRY
TITLE
hypothetical protein Rv3368c - Mycobacterium tuberculosis
(strain H37RV)
ORGANISM
#formal_name Mycobacterium tuberculosis
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
ACCESSIONS
REFERENCE
#authors
Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry
III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Hornaby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrell, B.G.

#journal
Nature (1998) 393:537-544
#title
Deciphering the biology of Mycobacterium tuberculosis from
the complete genome sequence.
#cross-references MUID:98295987
#accession
G70971
#status
preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-214 #label COL
#cross-references GB:AL023456; GB:AL123456; NID:g3342262; PID:e1202286;
PID:g2661647
#experimental_source strain H37RV
GENETICS
#gene
Rv3368c
SUMMARY
#length 214 #molecular-weight 23733 #checksum 2164
Query Match 30.0%; Score 6; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 200 AKRLPA 205
Qy 10 AKRLPA 15

RESULT 14
ENTRY
TITLE
deoxyurine kinase subunit homolog yaaF - Bacillus subtilis
ORGANISM
#formal_name Bacillus subtilis
DATE 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change
ACCESSIONS
REFERENCE
#authors
Ogasawara, N.; Nakai, S.; Yoshikawa, H.
#journal
DNA Res. (1994) 1:1-14
#title
Systematic sequencing of the 180 kilobase region of the
Bacillus subtilis chromosome containing the replication
origin.
#cross-references MUID:96051385
#accession
S66044
#status
preliminary; nucleic acid sequence not shown;
translation not shown

```

```

#molecule_type DNA
#residues 1-217 #label OGA
#cross-references EMBL:D26185; NID:g467326; PID:d1005792; PID:g467404
#note
the nucleotide sequence was submitted to the EMBL Data
Library, December 1993
REFERENCE
A69580
#authors
Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertoni, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Connor, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Dueterhoeft, A.;
Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goifeau, A.; Gollightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karanata, D.;
Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott,
A.M.; Presecan, E.; Fujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
#journal
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#title
#cross-references MUID:98044033
#accession
F69736
#status
preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-217 #label KUN
#cross-references GB:Z99104; GB:AL009126; NID:g2632267; PID:e1181947;
PID:g2632281
#experimental_source strain 168
GENETICS
#gene
yaaF
CLASSIFICATION
#superfamily Lactobacillus acidophilus deoxyadenosine kinase
SUMMARY
#length 217 #molecular-weight 25444 #checksum 714
Query Match 30.0%; Score 6; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 TLAKRL 33
Qy 8 TLAKRL 13

RESULT 15
ENTRY
TITLE
T03413 #type fragment
probable 1-aminocyclopropane-1-carboxylate synthase (EC
4.4.1.14) - rice (fragment)
ORGANISM
#formal_name Oryza sativa #common_name rice
DATE 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change

```



```

24-Mar-1999
ACCESSIONS T03413
REFERENCE Z14940
#authors van der Straeten, D.; Anuntalabhochai, S.; Zhou, Z.; Gielen,
J.; van Montagu, M.
#submission submitted to the EMBL Data Library, November 1993
#description Differential expression of three members of the ACC synthase
gene family in deepwater rice by submergence, wounding, and
hormonal treatments.
#accession T03413
##status preliminary; translated from GF/EMBL/DDBJ
##molecule_type mRNA
##residues 1-245 #label VAN
##cross-references EMBL:Z27242; NID:e970999; PID:e89105
##experimental_source cv. Plai Ngam, leaf
GENETICS
#gene ACS
KEYWORDS carbon-sulfur lyase
SUMMARY #length 245 #checksum 2188
Query Match 30.0%; Score 6; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.78e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 159 DRDTLA 164
Qy 5 DRDTLA 10
Search completed: Sat Aug 28 14:49:23 1999
Job time : 46 secs.

```

**This Page Blank (uspto)**

\*\*\*\*\*  
WISREH  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:46:24 1999; MasPar time 3.25 Seconds  
Tabular output not generated. 173.708 Million cell updates/sec

Title: >US-09-049-696-44  
Description: (1-20) from US09049696.pep  
Perfect Score: 20  
Sequence: 1 NSGSDRDTLAKRLPAAASGG 20

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 2.841; Variance 0.450; scale 6.313

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	8	40.0	1320	1	PUTA_SALTY	3.20e-03
2	7	35.0	153	1	MAL_MOUSE	1.97e-01
3	7	35.0	153	1	MAL_RAT	1.97e-01
4	7	35.0	153	1	MAL_CANFA	1.97e-01
5	6	30.0	113	1	RLA2_ALTAL	8.80e+00
6	6	30.0	123	1	GALA_HUMAN	8.80e+00
7	6	30.0	141	1	RL23_TYRCR	8.80e+00
8	6	30.0	157	1	FMPL_PSEAE	8.80e+00
9	6	30.0	168	1	ATPD_HUMAN	8.80e+00
10	6	30.0	172	1	UL45_HSV11	8.80e+00
11	6	30.0	172	1	UL45_HSV1K	8.80e+00
12	6	30.0	172	1	UL45_HSV1M	8.80e+00
13	6	30.0	196	1	CH19_DROGR	8.80e+00
14	6	30.0	209	1	NUGM_PAPRP	8.80e+00
15	6	30.0	217	1	YAAF_BACSU	8.80e+00
16	6	30.0	226	1	BASP_BOVIN	8.80e+00
17	6	30.0	247	1	YQFO_BACSU	8.80e+00
18	6	30.0	259	1	14-3-3-LIKE PROTEIN GF	8.80e+00
19	6	30.0	260	1	GTL1_CHICK	8.80e+00
20	6	30.0	265	1	CB22_WAIZE	8.80e+00
21	6	30.0	267	1	143C_ARATH	8.80e+00
22	6	30.0	297	1	AOK2_TOBAC	8.80e+00
23	6	30.0	297	1	NIFH_NOSCO	8.80e+00

60S ACIDIC RIBOSOMAL P 8.80e+00  
ELECTRON TRANSFER FLAV 8.80e+00  
GROWTH-ARREST-SPECIFIC 8.80e+00  
PRE-B-CELL LEUKEMIA TR 8.80e+00  
LIPASE MODULATOR PRECU 8.80e+00  
TRANSCRIPTION INITIAT 8.80e+00  
CYCLIN-DEPENDENT KINAS 8.80e+00  
GENERAL SECRETION PATH 8.80e+00  
LIPOPOLYSACCHARIDE 1,2 8.80e+00  
REGULATORY PROTEIN E2. 8.80e+00  
SEXUAL DIFFERENTIATION 8.80e+00  
PAIRED BOX POX-NEURO P 8.80e+00  
PRE-B-CELL LEUKEMIA TR 8.80e+00  
PRE-B-CELL LEUKEMIA TR 8.80e+00  
FUSARIC ACID RESISTANC 8.80e+00  
PRE-B-CELL LEUKEMIA TR 8.80e+00  
D-SERINE/D-ALANINE/GLY 8.80e+00  
VARIANT SURFACE GLYCOP 8.80e+00  
TUBBY PROTEIN. 8.80e+00  
FASCICLIN III PRECURSO 8.80e+00  
ZINC FINGER PROTEIN IA 8.80e+00  
RTN PROTEIN. 8.80e+00  
NITROGENASE MOLYBDENUM 8.80e+00  
PROBABLE THERMOSOME SU 8.80e+00  
HYPOTHETICAL 57.5 KD P 8.80e+00  
PROBABLE THERMOSOME SU 8.80e+00  
GLYCOPROTEIN E PRECURS 8.80e+00  
PHOSPHORIBOSYLAMINOIM 8.80e+00  
HEMAGGLUTININ-NEURAMIN 8.80e+00  
DMR-N9 PROTEIN. 8.80e+00  
HOMEOTIC-PROTEIN ORTHO 8.80e+00  
HYPOTHETICAL 77.5 KD P 8.80e+00  
TRANSFERIN-BINDING PR 8.80e+00  
PHENYLALANINE AMMONIA- 8.80e+00  
ATP-DEPENDENT CLP PROT 8.80e+00  
CELLULOSE SYNTHASE 93 8.80e+00  
DNA MISMATCH REPAIR PR 8.80e+00  
HYPOTHETICAL 171.5 KD 8.80e+00  
VITELLOGENIN I PRECURS 8.80e+00  
DNA-BINDING PROTEIN (A 2.57e+02  
HOMEBOX-PROTEIN CEH-9 2.57e+02  
PERIOD CLOCK PROTEIN ( 2.57e+02  
PERIOD CLOCK PROTEIN ( 2.57e+02  
30S RIBOSOMAL PROTEIN 2.57e+02  
HYPOTHETICAL 8.2 KD PR 2.57e+02  
STRUCTURAL PROTEIN VP2 2.57e+02  
CYTOCHROME C OXIDASE P 2.57e+02  
HYPOTHETICAL PROTEIN I 2.57e+02  
WIRIA PROTEIN. 2.57e+02  
PHOSPHOCARRIER PROTEIN 2.57e+02  
HYPOTHETICAL 10.8 KD P 2.57e+02  
60S ACIDIC RIBOSOMAL P 2.57e+02  
RIBOSOMAL PROTEIN L37 2.57e+02  
PHOSPHORIBOSYL-ATP CYC 2.57e+02  
HYPOTHETICAL 12.5 KD P 2.57e+02  
HYPOTHETICAL PROTEIN H 2.57e+02  
CAMP-REGULATED PHOSPHO 2.57e+02  
HYDROPHOBIN I PRECURSO 2.57e+02  
METHYLAMINE UTILIZATIO 2.57e+02  
VITELLINE MEMBRANE VM3 2.57e+02  
IG HEAVY CHAIN V-III R 2.57e+02  
50S RIBOSOMAL PROTEIN 2.57e+02  
MATING FACTOR ALPHA-2 2.57e+02  
SOMATOSTATIN I PRECURS 2.57e+02  
NADH-UBIQUINONE OXIDOR 2.57e+02  
PHOTOSYSTEM I REACTION 2.57e+02  
TRAD PROTEIN. 2.57e+02  
CHLOROPLAST 50S RIBOSO 2.57e+02  
RIBONUCLEASE P PROTEIN 2.57e+02  
ECTODERMAL DYSPLASIA P 2.57e+02

1 RLAO\_SOYBN 320  
1 ETFA\_RAT 333  
1 GAS1\_HUMAN 345  
1 PBX1\_MOUSE 347  
1 LINA\_PSEGL 353  
1 TF2B\_KLULA 357  
1 CD55\_HUMAN 367  
1 GSPL\_XANCP 373  
1 RFAK\_SALTY 381  
1 VE2\_HPV32 394  
1 ISP7\_SCHPO 397  
1 POXN\_DROME 425  
1 PBX2\_MOUSE 430  
1 PBX2\_HUMAN 430  
1 PBX1\_HUMAN 430  
1 FUSA\_HURCE 433  
1 PBX3\_HUMAN 434  
1 CYCA\_ECOLI 470  
1 VS16\_TRYBB 503  
1 TUB\_MOUSE 505  
1 FAS3\_DROME 508  
1 IAL\_HUMAN 510  
1 RTN\_ECOLI 518  
1 NIEK\_KLEPN 520  
1 THSB\_METH 538  
1 YU23\_MYCTH 542  
1 THSA\_METH 542  
1 VGLH\_HSVBS 575  
1 PU91\_YEAST 591  
1 PU92\_YEAST 592  
1 HEMA\_RINDR 609  
1 HEMA\_RINDK 609  
1 DMR9\_MOUSE 650  
1 HMOC\_DROME 671  
1 YK99\_YEAST 679  
1 TB21\_NEIME 711  
1 PALT\_RHOTO 716  
1 CLPA\_ECOLI 758  
1 ACSB\_ACEXY 802  
1 MUTS\_SALTY 854  
1 YB6A\_SCHPO 1522  
1 VITL\_CHICK 1512  
1 BVCP\_GVCL 58  
1 HM09\_CAEEL 63  
1 PER\_DROIM 63  
1 PER\_DROMS 63  
1 RS21\_HAEIN 70  
1 Y111\_NPVOP 72  
1 VP2\_SSV1 74  
1 COXI\_RAT 75  
1 YLDH\_ALCEU 77  
1 WIRA\_WHEAT 88  
1 PTHP\_BACST 88  
1 Y12K\_BPT4 95  
1 RLA2\_YEAST 106  
1 RL37\_TETTH 109  
1 HIS2\_AZOCH 110  
1 YDGC\_ECOLI 111  
1 YIFE\_HAEIN 111  
1 ARPP\_MOUSE 111  
1 HYPL\_AGABI 112  
1 MAUL\_METEX 114  
1 VTU3\_DROME 119  
1 HV3P\_HUMAN 119  
1 RL18\_SORBU 119  
1 MFA4\_YEAST 120  
1 SMS1\_LOPAM 121  
1 NB4M\_NEUCR 123  
1 PSAA\_SPIOL 125  
1 TRD5\_ECOLI 129  
1 RK12\_PORPU 129  
1 RNPA\_MICLU 132  
1 EDA\_HUMAN 135

97	5	25.0	136	1	Y05E_MYCTU	HYPOTHETICAL 15.4 KD P	2.57e+02	170	5	25.0	238	1	ARD1_YEAST	N-TERMINAL ACETYLTRANS	2.57e+02
98	5	25.0	137	1	RL23_YEAST	60S RIBOSOMAL PROTEIN	2.57e+02	171	5	25.0	239	1	CD8A_CANFA	T-CELL SURFACE GLYCOPR	2.57e+02
99	5	25.0	138	1	PSBU_SNP2	PHOTOSYSTEM II 12 KD E	2.57e+02	172	5	25.0	240	1	YGL2_HSVR	HYPOTHETICAL 23.9 KD P	2.57e+02
100	5	25.0	139	1	H2B_ENVI1	HISTONE H2B	2.57e+02	173	5	25.0	241	1	VG45_HSV1	HYPOTHETICAL GENE 45 P	2.57e+02
101	5	25.0	139	1	RL23_SCHPO	60S RIBOSOMAL PROTEIN	2.57e+02	174	5	25.0	242	1	Y956_HELPY	HYPOTHETICAL PROTEIN H	2.57e+02
102	5	25.0	140	1	RL23_DROME	60S RIBOSOMAL PROTEIN	2.57e+02	175	5	25.0	243	1	FRHG_METVO	COENZYME F420 HYDROGEN	2.57e+02
103	5	25.0	140	1	RL23_HUMAN	60S RIBOSOMAL PROTEIN	2.57e+02	176	5	25.0	243	1	HXB8_MOUSE	HOMEOBOX PROTEIN HOX-B	2.57e+02
104	5	25.0	141	1	CH16_DROSU	CHORION PROTEIN S16	2.57e+02	177	5	25.0	244	1	NDF3_RAT	NEUROGENIC DIFFERENTIA	2.57e+02
105	5	25.0	142	1	LACA_STAUA	GALACTOSE-6-PHOSPHATE	2.57e+02	178	5	25.0	245	1	YNF1_RHOCA	HYPOTHETICAL 27.7 KD P	2.57e+02
106	5	25.0	143	1	MERC_THIPE	MERCURIC RESISTANCE PR	2.57e+02	179	5	25.0	245	1	X355_HAEIN	HYPOTHETICAL ABC TRANS	2.57e+02
107	5	25.0	143	1	SYU_TORCA	SYNOCLEIN	2.57e+02	180	5	25.0	247	1	YH01_HAEIN	HYPOTHETICAL ABC TRANS	2.57e+02
108	5	25.0	145	1	Y1F1_YEAST	HYPOTHETICAL 15.9 KD P	2.57e+02	181	5	25.0	252	1	MODA_RHOCA	MOLYBDATE-BINDING PERI	2.57e+02
109	5	25.0	145	1	GFPG_ECOLI	PUTATIVE GENERAL SECRE	2.57e+02	182	5	25.0	252	1	DLTE_BACSU	DLTE PROTEIN	2.57e+02
110	5	25.0	149	1	H2B2_WHEAT	HISTONE H2B.2	2.57e+02	183	5	25.0	253	1	YZ11_MYCTU	HYPOTHETICAL 27.4 KD P	2.57e+02
111	5	25.0	150	1	VPS_BP2	TAIL COMPLETION PROTEI	2.57e+02	184	5	25.0	253	1	TRYD_DROME	TRYPSIN DELTA PRECURSO	2.57e+02
112	5	25.0	151	1	H2B1_MAIZE	HISTONE H2B.1	2.57e+02	185	5	25.0	253	1	TRYD_DROER	TRYPSIN DELTA PRECURSO	2.57e+02
113	5	25.0	152	1	H2B1_WHEAT	HISTONE H2B	2.57e+02	186	5	25.0	253	1	TRYB_DROER	TRYPSIN BETA PRECURSOR	2.57e+02
114	5	25.0	153	1	H2B3_MAIZE	HISTONE H2B.3	2.57e+02	187	5	25.0	253	1	TRYB_DROME	TRYPSIN GAMMA PRECURSO	2.57e+02
115	5	25.0	154	1	H2B5_MAIZE	HISTONE H2B	2.57e+02	188	5	25.0	255	1	YW24_MYCTU	HYPOTHETICAL 28.1 KD P	2.57e+02
116	5	25.0	155	1	Y1AL_ECOLI	CYTOSOLIC 17.5 KD P	2.57e+02	189	5	25.0	256	1	TRYA_DROME	TRYPSIN ALPHA PRECURSO	2.57e+02
117	5	25.0	155	1	C550_PARDE	CYTOSOLIC 17.5 KD P	2.57e+02	190	5	25.0	256	1	TRYA_DROER	TRYPSIN ALPHA PRECURSO	2.57e+02
118	5	25.0	155	1	RL30_PYRHO	50S RIBOSOMAL PROTEIN	2.57e+02	191	5	25.0	258	1	YURA_MYXSA	HYPOTHETICAL OXIDOREDU	2.57e+02
119	5	25.0	161	1	XCD1_BACSU	PHAGE-LIKE ELEMENT PBS	2.57e+02	192	5	25.0	260	1	XBPL_HUMAN	X BOX BINDING PROTEIN	2.57e+02
120	5	25.0	163	1	KU1_TENNO	PUPAL CUTICLE PROTEIN	2.57e+02	193	5	25.0	261	1	Y1GT_ECOLI	HYPOTHETICAL 27.8 KD P	2.57e+02
121	5	25.0	163	1	CU38_LOCMI	CUTICLE PROTEIN 38 (LM	2.57e+02	194	5	25.0	261	1	YU88_SYNY3	HISF PROTEIN (CYCLASE)	2.57e+02
122	5	25.0	165	1	TPX_MYCTU	PROBABLE THIOL PEROXID	2.57e+02	195	5	25.0	261	1	HIS6_SYNY3	HISF PROTEIN (CYCLASE)	2.57e+02
123	5	25.0	166	1	IL3_RAT	INTERLEUKIN-3 PRECURSO	2.57e+02	196	5	25.0	263	1	YXIS_SACER	HYPOTHETICAL 28.9 KD P	2.57e+02
124	5	25.0	166	1	RL21_ENTHI	60S RIBOSOMAL PROTEIN	2.57e+02	197	5	25.0	263	1	PM41_YEAST	MITOCHONDRIAL 60S RIBO	2.57e+02
125	5	25.0	173	1	AROL_ERWCH	SHKIMATE KINASE (EC 2	2.57e+02	198	5	25.0	265	1	DCOP_HANFA	OROTIDINE 5'-PHOSPHATE	2.57e+02
126	5	25.0	178	1	YFCB_ECOLI	HYPOTHETICAL 20.0 KD P	2.57e+02	199	5	25.0	265	1	EFID_XENLA	ELONGATION FACTOR 1-D	2.57e+02
127	5	25.0	179	1	UBCB_XENLA	UBIQUITIN-CONJUGATING	2.57e+02	200	5	25.0	265	1	MYO2_LYCES	MYO-INOSITOL-1(OR 4)-M	2.57e+02
128	5	25.0	180	1	RS5_SNP6	30S RIBOSOMAL PROTEIN	2.57e+02	201	5	25.0	266	1	Y1L1_MYCTU	HYPOTHETICAL 28.1 KD P	2.57e+02
129	5	25.0	181	1	PYER_BACSU	PYRIMIDINE OPERON REGU	2.57e+02	202	5	25.0	267	1	DCOP_HANAN	OROTIDINE 5'-PHOSPHATE	2.57e+02
130	5	25.0	185	1	YHHR_ECOLI	HYPOTHETICAL 19.8 KD P	2.57e+02	203	5	25.0	269	1	S3AD_SHIFL	STREPTOMYCIN 3'-ADENY	2.57e+02
131	5	25.0	186	1	YMA2_BRSAV	MATRIX GLYCOPROTEIN M2	2.57e+02	204	5	25.0	270	1	GPRL_YARLI	GLYCOXALATE PATHWAY RE	2.57e+02
132	5	25.0	186	1	YMA2_BRSVN	MATRIX GLYCOPROTEIN M2	2.57e+02	205	5	25.0	272	1	YCR3_ERWHE	HYPOTHETICAL 29.9 KD P	2.57e+02
133	5	25.0	190	1	CP28_HCMVA	28 KD STRUCTURAL PHOSP	2.57e+02	206	5	25.0	272	1	TYRO_STRAT	TYROSINASE (EC 1.14.18	2.57e+02
134	5	25.0	193	1	YCCB_BACSU	HYPOTHETICAL 22.3 KD P	2.57e+02	207	5	25.0	272	1	DNRI_STRPE	REGULATORY PROTEIN DNR	2.57e+02
135	5	25.0	194	1	LUXY_VIBFI	YELLOW FLUORESCENT PRO	2.57e+02	208	5	25.0	274	1	YOHM_ECOLI	HYPOTHETICAL 30.4 KD P	2.57e+02
136	5	25.0	194	1	YMA2_HRSVA	MATRIX GLYCOPROTEIN M2	2.57e+02	209	5	25.0	276	1	Y056_MYCPN	HYPOTHETICAL PROTEIN M	2.57e+02
137	5	25.0	197	1	KGUA_MOUSE	GUANYLATE KINASE (EC 2	2.57e+02	210	5	25.0	276	1	NO5Y_PSSET	MEMBRANE PROTEIN NOSY	2.57e+02
138	5	25.0	197	1	KGUA_PIG	GUANYLATE KINASE (EC 2	2.57e+02	211	5	25.0	277	1	DCOP_ASFNG	OROTIDINE 5'-PHOSPHATE	2.57e+02
139	5	25.0	197	1	KGUA_BOVIN	GUANYLATE KINASE (EC 2	2.57e+02	212	5	25.0	277	1	MYOP_MOUSE	MYO-INOSITOL-1(OR 4)-M	2.57e+02
140	5	25.0	198	1	RM12_HUMAN	MITOCHONDRIAL 60S RIBO	2.57e+02	213	5	25.0	277	1	BP42_STRAU	NON-HAEM CHLOROPEXID	2.57e+02
141	5	25.0	200	1	RAH1_ARATH	RAS-RELATED PROTEIN RH	2.57e+02	214	5	25.0	278	1	PRXC_STRAU	NON-HAEM CHLOROPEXID	2.57e+02
142	5	25.0	200	1	RHNI_NICPL	RAS-RELATED PROTEIN RH	2.57e+02	215	5	25.0	278	1	DCOP_ASFPU	OROTIDINE 5'-PHOSPHATE	2.57e+02
143	5	25.0	200	1	RAB5_TOBAC	RAS-RELATED PROTEIN RA	2.57e+02	216	5	25.0	279	1	PHEA_LACLA	PREPHENATE DEHYDRATASE	2.57e+02
144	5	25.0	201	1	RB1B_RAT	RAS-RELATED PROTEIN RA	2.57e+02	217	5	25.0	280	1	CCHL_CABEL	PROBABLE CYTOCHROME C-	2.57e+02
145	5	25.0	203	1	RM12_CRICR	MITOCHONDRIAL 60S RIBO	2.57e+02	218	5	25.0	281	1	RFSW_STRCO	RNA POLYMERASE SIGMA F	2.57e+02
146	5	25.0	204	1	IPF2_RABIT	PROTEIN PHOSPHATASE IN	2.57e+02	219	5	25.0	281	1	VG11_HSV1	HYPOTHETICAL GENE 11 Z	2.57e+02
147	5	25.0	204	1	PRAP_LEPBO	PUTATIVE PEPTIDASE PFA	2.57e+02	220	5	25.0	281	1	HS32_DICDI	32 KD HEAT SHOCK PROTE	2.57e+02
148	5	25.0	204	1	UREG_BACSE	UREASE ACCESSORY PROTE	2.57e+02	221	5	25.0	282	1	SP0J_BACSU	STAGE 0 SPORULATION PR	2.57e+02
149	5	25.0	205	1	YNS0_YEAST	HYPOTHETICAL 23.5 KD P	2.57e+02	222	5	25.0	284	1	YURL_BACSU	HYPOTHETICAL SUGAR KIN	2.57e+02
150	5	25.0	206	1	Y206_LAMBD	HYPOTHETICAL PROTEIN O	2.57e+02	223	5	25.0	284	1	KDSA_ECOLI	2-DEHYDRO-3-DEOXYPHOS	2.57e+02
151	5	25.0	207	1	CPVM_HUMAN	PEPTIDYL-PROLYL CIS-TR	2.57e+02	224	5	25.0	288	1	MYO2_HUMAN	MYO-INOSITOL-1(OR 4)-M	2.57e+02
152	5	25.0	209	1	CP23_CHICK	CP23 CORTICAL CYTOXKE	2.57e+02	225	5	25.0	291	1	RHO4_YEAST	RHO4 PROTEIN	2.57e+02
153	5	25.0	212	1	ALKH_HAEIN	PUTATIVE 4-HYDROXY-2-O	2.57e+02	226	5	25.0	291	1	USF2_RAT	UPSTREAM STIMULATORY F	2.57e+02
154	5	25.0	216	1	IFP2_ARATH	EUKARYOTIC TRANSLATION	2.57e+02	227	5	25.0	297	1	NLPD_PSEAF	LIPOPROTEIN NLPD/LPBP	2.57e+02
155	5	25.0	221	1	HIC_CHITE	HISTONE HIC	2.57e+02	228	5	25.0	298	1	RSF4_ARATH	40S RIBOSOMAL PROTEIN	2.57e+02
156	5	25.0	225	1	Y0DD_MYCTU	HYPOTHETICAL 24.0 KD P	2.57e+02	229	5	25.0	299	1	XJH9_YEAST	ACETYL-HYDROLASE (EC 3	2.57e+02
157	5	25.0	228	1	H13_GLYBA	HISTONE H1-III	2.57e+02	230	5	25.0	299	1	BAH_STRHY	DNASE I HOMOLOGOUS PRO	2.57e+02
158	5	25.0	229	1	KADH_TRIVA	ADENYLATE KINASE PRECU	2.57e+02	231	5	25.0	299	1	DHP1_HUMAN	60S RIBOSOMAL PROTEIN	2.57e+02
159	5	25.0	231	1	AP0Z_ECOLI	AQUAPORIN 2 (BACTERIAL	2.57e+02	232	5	25.0	299	1	RL22_DROME	60S RIBOSOMAL PROTEIN	2.57e+02
160	5	25.0	231	1	H1_CHITH	HISTONE H1	2.57e+02	233	5	25.0	300	1	NTRY_AZOB	NIROGEN REGULATION PR	2.57e+02
161	5	25.0	232	1	H12_GLYBA	HISTONE H1-II	2.57e+02	234	5	25.0	301	1	XADD_CHICK	HOMEOBOX PROTEIN HOX-D	2.57e+02
162	5	25.0	232	1	H1B_CHITE	HISTONE H1B	2.57e+02	235	5	25.0	305	1	SOHB_BUCAP	POSSIBLE PROTEASE SOHB	2.57e+02
163	5	25.0	232	1	H1A_CHITE	HISTONE H1A	2.57e+02	236	5	25.0	306	1	COQ4_YEAST	UBIQUINONE BIOSYNTHESI	2.57e+02
164	5	25.0	235	1	RS6_KLUMA	40S RIBOSOMAL PROTEIN	2.57e+02	237	5	25.0	307	1	IF2B_MALDO	EUKARYOTIC TRANSLATION	2.57e+02
165	5	25.0	236	1	RS6_YEAST	40S RIBOSOMAL PROTEIN	2.57e+02	238	5	25.0	307	1	COPE_CRIGR	COATOMER EPSILON SUBUN	2.57e+02
166	5	25.0	237	1	PIV6_ADEMI	MINOR CAPSID PROTEIN V	2.57e+02	239	5	25.0	308	1	Y04P_MYCTU	HYPOTHETICAL 33.0 KD P	2.57e+02
167	5	25.0	237	1	OCQ4_AGR16	OCOTOPINE TRANSPORT SYS	2.57e+02	240	5	25.0	309	1	TRUB_BACSU	TRNA PSEUDOURIDINE SYN	2.57e+02
168	5	25.0	237	1	RB23_MOUSE	RAS-RELATED PROTEIN RA	2.57e+02	241	5	25.0	312	1	IF2B_DROME	EUKARYOTIC TRANSLATION	2.57e+02
169	5	25.0	238	1	CHAC_ECOLI	CATION TRANSPORT PROTE	2.57e+02	242	5	25.0	316	1	YW34_MYCTU	HYPOTHETICAL TRANSCRIP	2.57e+02

243	5	25.0	317	1	K502	ACTCH	FRUIT PROTEIN PKIN1502	2.57e+02	316	5	25.0	376	1	KITH_HSV1K	THYMIDINE KINASE (EC 2	2.57e+02
244	5	25.0	318	1	HK25	MOUSE	HOMEOBOX PROTEIN NXX-2	2.57e+02	317	5	25.0	376	1	PSD4_MOUSE	26S. PROTEASOME REGULAT	2.57e+02
245	5	25.0	319	1	MTR1	RHOSH	MODIFICATION METHYLASE	2.57e+02	318	5	25.0	376	1	KITH_HSV11	THYMIDINE KINASE (EC 2	2.57e+02
246	5	25.0	321	1	YH28	YEAST	HYPOTHETICAL 36.5 KD P	2.57e+02	319	5	25.0	376	1	KITH_HSV1C	THYMIDINE KINASE (EC 2	2.57e+02
247	5	25.0	322	1	COAT	RSVT	COAT PROTEIN (NUCLEOCA	2.57e+02	320	5	25.0	377	1	SA18_MOUSE	TRANSCRIPTION FACTOR S	2.57e+02
248	5	25.0	323	1	SYK3	HABIN	HYPOTHETICAL LYSYL-TRN	2.57e+02	321	5	25.0	377	1	YA67_METTH	HYPOTHETICAL FACTOR M	2.57e+02
249	5	25.0	325	1	SYK3	ECOLI	HYPOTHETICAL LYSYL-TRN	2.57e+02	322	5	25.0	377	1	PSD4_HUMAN	26S. PROTEASOME REGULAT	2.57e+02
250	5	25.0	326	1	PBXA	HUMAN	PEROXISOME ASSEMBLY PR	2.57e+02	323	5	25.0	378	1	YHAD_HAEIN	HYPOTHETICAL PROTEIN H	2.57e+02
251	5	25.0	327	1	IDSA	METJA	BIFUNCTIONAL SHORT CHA	2.57e+02	324	5	25.0	378	1	PEL_PSESL	PECTATE LIASE. PRECURSO	2.57e+02
252	5	25.0	328	1	TEHA	HABIN	TELLURITE RESISTANCE P	2.57e+02	325	5	25.0	381	1	YHAD_ECOLI	HYPOTHETICAL 39.1 KD P	2.57e+02
253	5	25.0	328	1	VGH_BPG13		MINOR SPIKE PROTEIN (H	2.57e+02	326	5	25.0	381	1	YB2_ECOLI	HYPOTHETICAL 38.7 KD P	2.57e+02
254	5	25.0	328	1	VGH_BPHX		MINOR SPIKE PROTEIN (H	2.57e+02	327	5	25.0	382	1	YXAA_BACSU	HYPOTHETICAL 39.4 KD P	2.57e+02
255	5	25.0	330	1	TEHA	ECOLI	TELLURITE RESISTANCE P	2.57e+02	328	5	25.0	382	1	LPXB_ECOLI	LIPID-A-DISACCHARIDE S	2.57e+02
256	5	25.0	333	1	A85C	MCYLE	ANTIGEN 85-C PRECURSOR	2.57e+02	329	5	25.0	384	1	RN_DROME	GPASE ACTIVATING PROT	2.57e+02
257	5	25.0	333	1	CBBR_XANFL		RUBISCO OPERON TRANSCR	2.57e+02	330	5	25.0	389	1	ENAM_BOVIN	ENAMELIN (TUFTELIN) .	2.57e+02
258	5	25.0	333	1	G3P_STRAE		GLYCERALDEHYDE 3-PHOSP	2.57e+02	331	5	25.0	389	1	TRP1_MAIZE	TRYPTOPHAN SYNTHASE BE	2.57e+02
259	5	25.0	333	1	VG24_HSV11		HYPOTHETICAL GENE 24 P	2.57e+02	332	5	25.0	390	1	YFBE_ECOLI	HYPOTHETICAL 42.9 KD P	2.57e+02
260	5	25.0	335	1	HXD0	HUMAN	HOMEOBOX PROTEIN HOX-D	2.57e+02	333	5	25.0	390	1	LPXB_HAEIN	LIPID-A-DISACCHARIDE S	2.57e+02
261	5	25.0	336	1	FILA_MOUSE		FILAGRIN (FRAGMENT)	2.57e+02	334	5	25.0	391	1	NIRF_PSEST	NIRF PROTEIN	2.57e+02
262	5	25.0	338	1	G3P_SCHNA		GLYCERALDEHYDE 3-PHOSP	2.57e+02	335	5	25.0	394	1	TYRB_PARDE	AROMATIC-AMINO-ACID AM	2.57e+02
263	5	25.0	339	1	HXD9_MOUSE		HOMEOBOX PROTEIN HOX-D	2.57e+02	336	5	25.0	394	1	TCR4_SALOR	TETRACYCLINE RESISTANC	2.57e+02
264	5	25.0	339	1	MOXR_PARDE		MOXR PROTEIN.	2.57e+02	337	5	25.0	395	1	YT44_STRFR	HYPOTHETICAL 44.4 KD P	2.57e+02
265	5	25.0	340	1	YXJH_BACSU		HYPOTHETICAL 38.3 KD P	2.57e+02	338	5	25.0	395	1	SDC_DROME	SYNDECAN PRECURSOR.	2.57e+02
266	5	25.0	340	1	ALGP_PSEAE		TRANSCRIPTIONAL REGULA	2.57e+02	339	5	25.0	395	1	KINE_RAT	MEVALONATE KINASE (EC	2.57e+02
267	5	25.0	341	1	YCFO_ECOLI		HYPOTHETICAL 37.6 KD P	2.57e+02	340	5	25.0	396	1	HMPA_ECOLI	FLAVONOHOPROTEIN (HAEM	2.57e+02
268	5	25.0	341	1	MOXR_MEYEX		MOXR PROTEIN (MXAR PRO	2.57e+02	341	5	25.0	397	1	METK_STAOU	S-ADENOSYLMETHIONINE S	2.57e+02
269	5	25.0	342	1	ISIA_SYNP7		IRON-STRESS INDUCED CH	2.57e+02	342	5	25.0	398	1	VEZ2_HP42	REGULATORY PROTEIN E2.	2.57e+02
270	5	25.0	344	1	NLPB_ECOLI		LIPOPROTEIN-34 PRECURS	2.57e+02	343	5	25.0	399	1	PRTZ_HORVU	PROTEIN 2 (24) (MAJOR	2.57e+02
271	5	25.0	344	1	TOB4_HUMAN		TOB4 PROTEIN.	2.57e+02	344	5	25.0	402	1	OPS4_CANAL	OPAQUE-PHASE-SPECIFIC	2.57e+02
272	5	25.0	346	1	BIAR_SHEEP		BETA-1 ADRENERGIC RECE	2.57e+02	345	5	25.0	403	1	TYRP_ECOLI	TYROSINE-SPECIFIC TRAN	2.57e+02
273	5	25.0	346	1	CCHL_NEUCR		CYTOCHROME C HEME LYAS	2.57e+02	346	5	25.0	405	1	DCP2_PEA	PYRUVATE DECARBOXYLASE	2.57e+02
274	5	25.0	346	1	USF2_MOUSE		UPSTREAM STIMULATORY F	2.57e+02	347	5	25.0	406	1	DYNC_HUMAN	DYNACTIN, 50 KD ISOFOR	2.57e+02
275	5	25.0	346	1	USF2_MOUSE		UPSTREAM STIMULATORY F	2.57e+02	348	5	25.0	406	1	OTE_DROME	OTEFIN.	2.57e+02
276	5	25.0	348	1	TRPD_CORGL		ANTHRANILATE PHOSPHORI	2.57e+02	349	5	25.0	407	1	YX23_MYCTU	HYPOTHETICAL 41.5 KD P	2.57e+02
277	5	25.0	348	1	TRPD_BRELA		ANTHRANILATE PHOSPHORI	2.57e+02	350	5	25.0	408	1	AK2_BACSU	ASPARTOKINASE 2 ALPHA	2.57e+02
278	5	25.0	349	1	SOHB_ECOLI		POSSIBLE PROTEASE SOHB	2.57e+02	351	5	25.0	408	1	SEPR_THESP	EXTRACELLULAR SERINE P	2.57e+02
279	5	25.0	349	1	GUB_FISBU		BETA-GLUCANASE PRECURS	2.57e+02	352	5	25.0	408	1	HELK_PROTEIN.	HELK PROTEIN.	2.57e+02
280	5	25.0	351	1	CBIG_SALTY		CBIG PROTEIN..	2.57e+02	353	5	25.0	410	1	HFLK_HAEIN	NEUROSERPIN PRECURSOR	2.57e+02
281	5	25.0	352	1	REP3_THIFE		DNA PRIMASE (EC 2.7.7.	2.57e+02	354	5	25.0	410	1	NEUS_MOUSE	ELONGATION FACTOR TU (	2.57e+02
282	5	25.0	352	1	KITH_HSVEB		THYMIDINE KINASE (EC 2	2.57e+02	355	5	25.0	410	1	EFTU_SPIPL	NEUROSERPIN PRECURSOR	2.57e+02
283	5	25.0	352	1	KITH_HSVEA		THYMIDINE KINASE (EC 2	2.57e+02	356	5	25.0	410	1	NEUS_CHICK	NEUROSERPIN PRECURSOR	2.57e+02
284	5	25.0	353	1	SOHB_HAEIN		POSSIBLE PROTEASE SOHB	2.57e+02	357	5	25.0	413	1	CAOC_RAT	CALDESQUESTRIN, CARDIAC	2.57e+02
285	5	25.0	353	1	XYNA_PENCH		ENDO-1,4-BETA-XYLANASE	2.57e+02	358	5	25.0	414	1	YGJU_HAEIN	HYPOTHETICAL SYMPORTER	2.57e+02
286	5	25.0	354	1	TH41_MAIZE		THIAMINE BIOSYNTHETIC	2.57e+02	359	5	25.0	415	1	SPHS_SYNP7	SENSOR PROTEIN SPHS (E	2.57e+02
287	5	25.0	354	1	OTC_MOUSE		ORNITHINE CARBAMOYLTRA	2.57e+02	360	5	25.0	415	1	TRAI_MYCTU	TRANSPOSASE FOR INSERT	2.57e+02
288	5	25.0	354	1	OTC_MOUSE		ORNITHINE CARBAMOYLTRA	2.57e+02	361	5	25.0	416	1	MYC_AVIOK	MYC TRANSFORMING PROTE	2.57e+02
289	5	25.0	357	1	MTDC_DROME		NAD-DEPENDENT METHYLEN	2.57e+02	362	5	25.0	416	1	YEIM_ECOLI	HYPOTHETICAL 43.4 KD P	2.57e+02
290	5	25.0	359	1	ALRI_ECOLI		ALANINE RACEMASE, BIOS	2.57e+02	363	5	25.0	416	1	PI2R_RAT	PROSTACYCLIN RECEPTOR	2.57e+02
291	5	25.0	359	1	ALRI_SALTY		ALANINE RACEMASE, BIOS	2.57e+02	364	5	25.0	416	1	CAQC_MOUSE	CALDESQUESTRIN, CARDIAC	2.57e+02
292	5	25.0	360	1	ALR2_HAEIN		ALANINE RACEMASE, BIOS	2.57e+02	365	5	25.0	416	1	HXD3_HUMAN	HOMEOBOX PROTEIN HOX-D	2.57e+02
293	5	25.0	361	1	ELY2_BACSP		THERMOSTABLE ALKALINE	2.57e+02	366	5	25.0	416	1	MYC_CHICK	MYC PROTO-ONCOGENE PRO	2.57e+02
294	5	25.0	361	1	FDEH_PSEPU		5'-EXO-ALCOHOL DEHYDROG	2.57e+02	367	5	25.0	417	1	PROA_SERMA	GAMMA-GLUTAMYL PHOSPHA	2.57e+02
295	5	25.0	361	1	BRNA_MOUSE		BRAIN-SPECIFIC HOMEOBO	2.57e+02	368	5	25.0	418	1	DCP1_TOBAC	PYRUVATE DECARBOXYLASE	2.57e+02
296	5	25.0	361	1	BRNA_MOUSE		BRAIN-SPECIFIC HOMEOBO	2.57e+02	369	5	25.0	419	1	Y092_HAEIN	HYPOTHETICAL PROTEIN H	2.57e+02
297	5	25.0	362	1	YBDL_ECOLI		HYPOTHETICAL OXIDOREDU	2.57e+02	370	5	25.0	419	1	GYRE_VIBPA	DNA GYRASE SUBUNIT B (	2.57e+02
298	5	25.0	363	1	SSR5_HUMAN		SOMATOSTATIN RECEPTOR	2.57e+02	371	5	25.0	420	1	GSC_DROME	HOMEOBOX PROTEIN REGUL	2.57e+02
299	5	25.0	364	1	PGLR_COCCA		POLYGALACTURONASE PREC	2.57e+02	372	5	25.0	420	1	PSD4_SCHMA	26S. PROTEASOME REGOAT	2.57e+02
300	5	25.0	364	1	KITH_ILTVT		THYMIDINE KINASE (EC 2	2.57e+02	373	5	25.0	421	1	CU4A_HUMAN	CULLIN HOMOLOG 4A (CUL	2.57e+02
301	5	25.0	366	1	GDF3_MOUSE		GROWTH/DIFFERENTIATION	2.57e+02	374	5	25.0	421	1	MYC_AVIM2	MYC TRANSFORMING PROTE	2.57e+02
302	5	25.0	367	1	AROG_CORGL		PHOSPHO-2-DEHYDRO-3-DE	2.57e+02	375	5	25.0	422	1	MYC_AVIM2	MYC TRANSFORMING PROTE	2.57e+02
303	5	25.0	368	1	DHAS_ECOLI		ASPARTATE-SEMIALDEHYDE	2.57e+02	376	5	25.0	422	1	MYC_AVIMC	MYC TRANSFORMING PROTE	2.57e+02
304	5	25.0	368	1	DHAS_SALTY		ASPARTATE-SEMIALDEHYDE	2.57e+02	377	5	25.0	423	1	MYC_AVIME	MYC TRANSFORMING PROTE	2.57e+02
305	5	25.0	370	1	ATPG_OOOSI		ATP SYNTHASE GAMMA CHA	2.57e+02	378	5	25.0	423	1	CHIL_APHAL	CHITINASE 1 PRECURSOR	2.57e+02
306	5	25.0	372	1	HXA2_MOUSE		HOMEOBOX PROTEIN HOX-A	2.57e+02	379	5	25.0	426	1	YXK4_YEAST	HYPOTHETICAL 47.8 KD P	2.57e+02
307	5	25.0	372	1	HXA2_RAT		HOMEOBOX PROTEIN HOX-A	2.57e+02	380	5	25.0	427	1	GSBD_DROME	GOOSEBERRY DISTAL PROT	2.57e+02
308	5	25.0	373	1	EGON_DROME		EMBRYONIC GONAD PROTEI	2.57e+02	381	5	25.0	437	1	AFLR_ASPFL	AFATOXIN BIOSYNTHESIS	2.57e+02
309	5	25.0	373	1	TYRA_ERWHE		CHORISMATE MUTASE (EC	2.57e+02	382	5	25.0	438	1	GCDD_HUMAN	GLUTARYL-COA DEHYDROGE	2.57e+02
310	5	25.0	373	1	TYRA_ECOLI		CHORISMATE MUTASE (EC	2.57e+02	383	5	25.0	439	1	TEF5_MOUSE	TRANSCRIPTIONAL ENHANC	2.57e+02
311	5	25.0	374	1	RF2_MCYLE		PEPTIDE CHAIN RELEASE	2.57e+02	384	5	25.0	441	1	IF5_CAEEL	EUKARYOTIC TRANSLATION	2.57e+02
312	5	25.0	375	1	SOX3_MOUSE		TRANSCRIPTION FACTOR S	2.57e+02	385	5	25.0	441	1	RRD2_NOTVI	RETINOIC ACID RECEPTOR	2.57e+02
313	5	25.0	376	1	DHAS_BORPE		ASPARTATE-SEMIALDEHYDE	2.57e+02	386	5	25.0	441	1	DNB2_ADEGI	EARLY A2A DNA-BINDING	2.57e+02
314	5	25.0	376	1	KITH_HSVIS		THYMIDINE KINASE (EC 2	2.57e+02	387	5	25.0	442	1	CELF_PROBABLE	6-PHOSPHO-BET	2.57e+02
315	5	25.0	376	1	KITH_HSVIE		THYMIDINE KINASE (EC 2	2.57e+02	388	5	25.0	442	1	FTSA_RHIME	CELL DIVISION PROTEIN	2.57e+02

389	5	25.0	442	1	RRG2_XENLA	RETINOIC ACID RECEPTOR	2.57e+02	462	5	25.0	492	1	MENE_STAAR	O-SUCCINYL BENZOIC ACID	2.57e+02
390	5	25.0	443	1	SOX3_HUMAN	TRANSCRIPTION FACTOR S	2.57e+02	463	5	25.0	493	1	1A11_CUCPE	1-AMINOCYCLOPROPANE-1-	2.57e+02
391	5	25.0	443	1	RRG2_HUMAN	RETINOIC ACID RECEPTOR	2.57e+02	464	5	25.0	493	1	1A11_CUCMA	1-AMINOCYCLOPROPANE-1-	2.57e+02
392	5	25.0	443	1	TRP2_MAIZE	TRYPTOPHAN SYNTHASE BE	2.57e+02	465	5	25.0	494	1	G6PD_HAETH	GLUCOSE-6-PHOSPHATE 1-	2.57e+02
393	5	25.0	444	1	YHAU_ECOLI	HYPOTHETICAL 49.0 KD P	2.57e+02	466	5	25.0	494	1	1A12_CUCPE	1-AMINOCYCLOPROPANE-1-	2.57e+02
394	5	25.0	444	1	T1SD_ECOLI	TYPE I RESTRICTION ENZ	2.57e+02	467	5	25.0	494	1	G6PD_ACTAC	GLUCOSE-6-PHOSPHATE 1-	2.57e+02
395	5	25.0	444	1	AFIR_ASPPA	AFLATOXIN BIOSYNTHESIS	2.57e+02	468	5	25.0	494	1	DCE2_ARATH	GLUTAMATE DECARBOXYLAS	2.57e+02
396	5	25.0	446	1	YHE2_PSEAE	HYPOTHETICAL 50.0 KD P	2.57e+02	469	5	25.0	500	1	DCE1_PETHY	GLUTAMATE DECARBOXYLAS	2.57e+02
397	5	25.0	446	1	GABT_MYLE	4-AMINOBUTYRATE AMINOT	2.57e+02	470	5	25.0	500	1	DCE1_ARATH	GLUTAMATE DECARBOXYLAS	2.57e+02
398	5	25.0	446	1	SYG_MYCGE	GLYCYL-TRNA SYNTHASE	2.57e+02	471	5	25.0	502	1	DCE1_LYCES	GLUTAMATE DECARBOXYLAS	2.57e+02
399	5	25.0	446	1	PFES_PSEAE	SENSOR PROTEIN PFES (E	2.57e+02	472	5	25.0	505	1	RDL1_NOTVI	RETINOIC ACID RECEPTOR	2.57e+02
400	5	25.0	446	1	PIUR_PSEAE	TYPE 4 FIMBRIAE EXPRES	2.57e+02	473	5	25.0	507	1	YG46_YEAST	HYPOTHETICAL 52.8 KD P	2.57e+02
401	5	25.0	447	1	RRG2_MOUSE	RETINOIC ACID RECEPTOR	2.57e+02	474	5	25.0	508	1	TDH_HUMAN	DNA NUCLEOTIDYLKEXTRAN	2.57e+02
402	5	25.0	448	1	RRB2_CHICK	RETINOIC ACID RECEPTOR	2.57e+02	475	5	25.0	509	1	DLDH_HUMAN	DIHYDROLIPOAMIDE DEHYD	2.57e+02
403	5	25.0	448	1	RRB2_HUMAN	RETINOIC ACID RECEPTOR	2.57e+02	476	5	25.0	509	1	DLDH_PIG	DIHYDROLIPOAMIDE DEHYD	2.57e+02
404	5	25.0	448	1	HGD_EWENI	HOMOGENITISATE 1, 2-DIOX	2.57e+02	477	5	25.0	509	1	ATPA_RHOCA	ATP SYNTHASE ALPHA CHA	2.57e+02
405	5	25.0	449	1	AAAT3_ARATH	ASPARTATE AMINOTRANSE	2.57e+02	478	5	25.0	510	1	GLPD_PSEAE	GLYCEROL-3-PHOSPHATE D	2.57e+02
406	5	25.0	451	1	ARGD_AUNGL	ACETYLORNITHINE AMINOT	2.57e+02	479	5	25.0	511	1	COT4_BACSU	SPORE COAT PROTEIN A.	2.57e+02
407	5	25.0	452	1	GLPT_ECOLI	GLYCEROL-3-PHOSPHATE T	2.57e+02	480	5	25.0	513	1	BMP6_HUMAN	BONE MORPHOGENETIC PRO	2.57e+02
408	5	25.0	454	1	RRG1_HUMAN	RETINOIC ACID RECEPTOR	2.57e+02	481	5	25.0	513	1	TYRR_ECOLI	TRANSCRIPTIONAL REGULA	2.57e+02
409	5	25.0	454	1	NOLI_RHIME	MODULATION PROTEIN NOL	2.57e+02	482	5	25.0	516	1	Y195_MYCLE	HYPOTHETICAL 56.0 KD P	2.57e+02
410	5	25.0	455	1	Y0D2_MYCTU	HYPOTHETICAL 48.1 KD P	2.57e+02	483	5	25.0	516	1	ACHD_BOVIN	ACETYLCHOLINE RECEPTOR	2.57e+02
411	5	25.0	455	1	RRB1_CHICK	RETINOIC ACID RECEPTOR	2.57e+02	484	5	25.0	517	1	ACHD_HUMAN	ACETYLCHOLINE RECEPTOR	2.57e+02
412	5	25.0	456	1	Y0Z7_MYCTU	HYPOTHETICAL 49.6 KD P	2.57e+02	485	5	25.0	517	1	ACHD_RAT	ACETYLCHOLINE RECEPTOR	2.57e+02
413	5	25.0	457	1	LEF4_NPVOP	LATE EXPRESSION FACTOR	2.57e+02	486	5	25.0	518	1	TTD_MONDO	DNA NUCLEOTIDYLKEXTRAN	2.57e+02
414	5	25.0	458	1	RRG1_MOUSE	RETINOIC ACID RECEPTOR	2.57e+02	487	5	25.0	519	1	MOD3_MOUSE	MODIFIER 3 PROTEIN (M3	2.57e+02
415	5	25.0	458	1	TRKA_VIBAL	TRK SYSTEM POTASSIUM U	2.57e+02	488	5	25.0	520	1	ACHD_MOUSE	ACETYLCHOLINE RECEPTOR	2.57e+02
416	5	25.0	459	1	XDUL_YEAST	XDJ1 PROTEIN	2.57e+02	489	5	25.0	521	1	PMNR_NEIGO	PEPTIDE METHIONINE SUL	2.57e+02
417	5	25.0	461	1	PSBC_CYAPA	PHOTOSYSTEM II 44 KD R	2.57e+02	490	5	25.0	525	1	SYH_CAEEL	HISTIDYL-TRNA SYNTHETA	2.57e+02
418	5	25.0	462	1	GSA_MYCTU	INDOLEACETAMIDE HYDROL	2.57e+02	491	5	25.0	525	1	NAB2_YEAST	NUCLEAR POLYADENYLATED	2.57e+02
419	5	25.0	462	1	HYIN_AGRVI	INDOLEACETAMIDE HYDROL	2.57e+02	492	5	25.0	527	1	PPOX_PROPR	PROTOPORPHYRINOGEN OXI	2.57e+02
420	5	25.0	463	1	PHYB_EWENI	3-PHYTASE B PRECURSOR	2.57e+02	493	5	25.0	527	1	FSUB_SERMA	IRON (III)-TRANSPORT S	2.57e+02
421	5	25.0	464	1	T1SK_ECOLI	TYPE I RESTRICTION ENZ	2.57e+02	494	5	25.0	528	1	HS60_AJECA	HEAT SHOCK PROTEIN 60	2.57e+02
422	5	25.0	465	1	SELA_DESBA	L-SERYL-TRNA(SEC) SELE	2.57e+02	495	5	25.0	529	1	YGNA_ECOLI	HYPOTHETICAL SIGMA-54-	2.57e+02
423	5	25.0	465	1	SPF4_CAEEL	INTEGRAL MEMBRANE PROT	2.57e+02	496	5	25.0	529	1	TDI_MOUSE	DNA NUCLEOTIDYLKEXTRAN	2.57e+02
424	5	25.0	466	1	ZIC3_MOUSE	ZINC FINGER PROTEIN ZI	2.57e+02	497	5	25.0	531	1	UDI6_HUMAN	UDP-GLUCURONOSYLTRANSF	2.57e+02
425	5	25.0	466	1	DCEB_ECOLI	GLUTAMATE DECARBOXYLAS	2.57e+02	498	5	25.0	532	1	YABM_BACSU	HYPOTHETICAL 57.4 KD P	2.57e+02
426	5	25.0	466	1	DCEA_ECOLI	GLUTAMATE DECARBOXYLAS	2.57e+02	499	5	25.0	533	1	SYK_METMP	HYPOHETICAL 57.4 KD P	2.57e+02
427	5	25.0	466	1	TRPB_CAMAC	TRYPTOPHAN SYNTHASE BE	2.57e+02	500	5	25.0	534	1	FMO2_MACMU	DIMETHYLANILINE MONOOX	2.57e+02
428	5	25.0	466	1	HYIN_AGRRA	INDOLEACETAMIDE HYDROL	2.57e+02	501	5	25.0	536	1	SKIP_HUMAN	NUCLEAR PROTEIN SKIP (	2.57e+02
429	5	25.0	467	1	ZIC3_HUMAN	ZINC FINGER PROTEIN ZI	2.57e+02	502	5	25.0	536	1	AROC_TOXGO	CHORISMATE SYNTHASE (E	2.57e+02
430	5	25.0	467	1	HYIN_AGRTA	INDOLEACETAMIDE HYDROL	2.57e+02	503	5	25.0	542	1	THS_METJA	NUCLEOCAPSID SUBUNIT	2.57e+02
431	5	25.0	467	1	HYIN_AGRT3	INDOLEACETAMIDE HYDROL	2.57e+02	504	5	25.0	543	1	NCAP_SV41	NUCLEOCAPSID PROTEIN	2.57e+02
432	5	25.0	468	1	BLAR_PIG	BETA-1 ADRENERGIC RECE	2.57e+02	505	5	25.0	544	1	SYK_METTH	LYSYL-TRNA SYNTHETASE	2.57e+02
433	5	25.0	469	1	T1S_SALTY	TYPE I RESTRICTION ENZ	2.57e+02	506	5	25.0	546	1	ODP2_PSEAE	DIHYDROLIPOAMIDE ACETY	2.57e+02
434	5	25.0	470	1	TRP1_ARATH	TRYPTOPHAN SYNTHASE BE	2.57e+02	507	5	25.0	548	1	Y24C_MYCTU	HYPOTHETICAL 57.9 KD P	2.57e+02
435	5	25.0	474	1	ILBA_ECOLI	HYPOTHETICAL 53.7 KD P	2.57e+02	508	5	25.0	548	1	GLL_YEAST	TRIGLYCERIDE LIPASE-CH	2.57e+02
436	5	25.0	474	1	T1SB_ECOLI	TYPE I RESTRICTION ENZ	2.57e+02	509	5	25.0	550	1	CCF_DROME	CENTROSOMAL AND CHROMO	2.57e+02
437	5	25.0	474	1	UCR2_EUGGR	UBIQUINOL-CYTOCHROME-C	2.57e+02	510	5	25.0	551	1	ATPA_NEUCR	ATP SYNTHASE ALPHA CHA	2.57e+02
438	5	25.0	475	1	TRP2_ARATH	TRYPTOPHAN SYNTHASE BE	2.57e+02	511	5	25.0	553	1	MCRA_METVA	METHYL-COENZYME M REDU	2.57e+02
439	5	25.0	475	1	IDH_SNY3	ISOCITRATE DEHYDROGENA	2.57e+02	512	5	25.0	553	1	ODP2_ALCEU	DIHYDROLIPOAMIDE ACETY	2.57e+02
440	5	25.0	475	1	EVX2_MOUSE	HOMSOBOX EVEN-SKIPPED	2.57e+02	513	5	25.0	555	1	MURA_METFE	METHYL-COENZYME M REDU	2.57e+02
441	5	25.0	476	1	IA14_LYCES	1-AMINOCYCLOPROPANE-1-	2.57e+02	514	5	25.0	556	1	PID_STRAT	PHOSPHOLIPASE D PRECUR	2.57e+02
442	5	25.0	476	1	RRGL_XENLA	RETINOIC ACID RECEPTOR	2.57e+02	515	5	25.0	557	1	TKT2_HUMAN	TRANSGUTOLASE 2 (EC 2	2.57e+02
443	5	25.0	477	1	FES_FSVST	TYROSINE-PROTEIN KINAS	2.57e+02	516	5	25.0	558	1	DHE4_HUMAN	GLUTAMATE DEHYDROGENAS	2.57e+02
444	5	25.0	477	1	RBL_LACSA	RIBULOSE BISPHOSPHATE	2.57e+02	517	5	25.0	559	1	PHAA_PSEOL	POLY(3-HYDROXYALKANOAT	2.57e+02
445	5	25.0	477	1	PTGB_ECOLI	PTS SYSTEM, GLUCOSE-SP	2.57e+02	518	5	25.0	560	1	PHAC_PSEOL	POLY(3-HYDROXYALKANOAT	2.57e+02
446	5	25.0	477	1	PTGB_SALTY	PTS SYSTEM, GLUCOSE-SP	2.57e+02	519	5	25.0	560	1	GPV_HUMAN	PLATELET GLYCOPROTEIN	2.57e+02
447	5	25.0	477	1	MAZ_HUMAN	MYC-ASSOCIATED ZINC FI	2.57e+02	520	5	25.0	561	1	AIAD_RAT	ALPHA-1D ADRENERGIC RE	2.57e+02
448	5	25.0	478	1	LXCI_PHOLE	ACYL-COA REDUCTASE (EC	2.57e+02	521	5	25.0	561	1	BCDH_RHOCA	MAGNESIUM-CHLATELASE 60	2.57e+02
449	5	25.0	478	1	BMB3_HUMAN	BONE MORPHOGENETIC PRO	2.57e+02	522	5	25.0	562	1	AIAD_MOUSE	ALPHA-1D ADRENERGIC RE	2.57e+02
450	5	25.0	478	1	ERG_CHICK	TRANSCRIPTIONAL REGULA	2.57e+02	523	5	25.0	564	1	OXC_ECOLI	PROBABLE OXALYL-COA DE	2.57e+02
451	5	25.0	480	1	PY95_BOVIN	URIDINE 5'-MONOPHOSPHA	2.57e+02	524	5	25.0	565	1	TREA_ECOLI	PERIPLASMIC TREHALASE	2.57e+02
452	5	25.0	482	1	Y269_TREPA	HYPOTHETICAL PROTEIN T	2.57e+02	525	5	25.0	567	1	GPV_RAT	PLATELET GLYCOPROTEIN	2.57e+02
453	5	25.0	482	1	RRB_MOUSE	RETINOIC ACID RECEPTOR	2.57e+02	526	5	25.0	567	1	GPV_MOUSE	PLATELET GLYCOPROTEIN	2.57e+02
454	5	25.0	483	1	TR2_HUMAN	ORPHAN RECEPTOR TR2	2.57e+02	527	5	25.0	568	1	RS1_RHIME	30S RIBOSOMAL PROTEIN	2.57e+02
455	5	25.0	484	1	YL55_CAEEL	HYPOTHETICAL 55.1 KD P	2.57e+02	528	5	25.0	569	1	UVRA_VITST	EXCINUCLEASE ABC SUBUN	2.57e+02
456	5	25.0	485	1	G6PD_AYMMO	GLUCOSE-6-PHOSPHATE 1-	2.57e+02	529	5	25.0	571	1	KNLC_LOLPE	KINESIN LIGHT CHAIN (K	2.57e+02
457	5	25.0	485	1	NWRB_RAT	NEURONAL OLFACTOMEDIN-	2.57e+02	530	5	25.0	571	1	LCFH_MYCTU	PUTATIVE LONG-CHAIN-F	2.57e+02
458	5	25.0	485	1	RT16_MTXXA	RNA-DIRECTED DNA POLYM	2.57e+02	531	5	25.0	581	1	CLK9_HUMAN	VOLTAGE-GATED POTASSIU	2.57e+02
459	5	25.0	486	1	YDF1_ECOLI	HYPOTHETICAL 53.7 KD P	2.57e+02	532	5	25.0	582	1	MNT_HUMAN	MUSC BINDING PROTEIN MN	2.57e+02
460	5	25.0	491	1	Y429_SCHPO	HYPOTHETICAL 55.1 KD P	2.57e+02	533	5	25.0	589	1	ACM3_RAT	MUSCARINIC ACETYLCHOLI	2.57e+02
461	5	25.0	491	1	1A1C_TOBAC	1-AMINOCYCLOPROPANE-1-	2.57e+02	534	5	25.0	589	1	VP40_SCMVC	CAPSID PROTEIN P40 [CO	2.57e+02

535	5	25.0	589	1	PHBC_ALCEU	POLY-BETA-HYDROXYBUTYR	2.57e+02	608	5	25.0	702	1	AT1L_VARV	81 KD A-TYPE INCLUSION	2.57e+02
536	5	25.0	590	1	STUA_EMENI	CELL PATTERN FORMATION	2.57e+02	609	5	25.0	704	1	HS90_SCHPO	HEAT SHOCK PROTEIN 90	2.57e+02
537	5	25.0	591	1	ACM3_HUMAN	MUSCARINIC ACETYLCHOLI	2.57e+02	610	5	25.0	707	1	DREB_RAT	DREBRIN A	2.57e+02
538	5	25.0	590	1	ACM3_BOVIN	MUSCARINIC ACETYLCHOLI	2.57e+02	611	5	25.0	707	1	TRP_YEAST	TRYPTOPHAN SYNTHASE (E	2.57e+02
539	5	25.0	591	1	ACM3_PIG	MUSCARINIC ACETYLCHOLI	2.57e+02	612	5	25.0	711	1	PO22_POPJA	RETROVIRUS-RELATED POL	2.57e+02
540	5	25.0	591	1	PUR3_HUMAN	PHOSPHORIBOSYLAMINOIM	2.57e+02	613	5	25.0	716	1	ARA_DROME	HOMEOBOX PROTEIN ARAUC	2.57e+02
541	5	25.0	593	1	YJX2_YEAST	HYPOTHETICAL 67.0 KD P	2.57e+02	614	5	25.0	718	1	ICAL_RABIT	CALPAIN INHIBITOR PREC	2.57e+02
542	5	25.0	593	1	PUR3_CHICK	HYPOTHETICAL 67.0 KD P	2.57e+02	615	5	25.0	724	1	PRIP_PRRIF	PROBABLE PROCESSING AN	2.57e+02
543	5	25.0	594	1	RYK_MOUSE	PHOSPHORIBOSYLAMINOIM	2.57e+02	616	5	25.0	724	1	TAP1_MOUSE	PROBABLE PROCESSING AN	2.57e+02
544	5	25.0	594	1	SYFA_YEAST	TYROSINE-PROTEIN KINAS	2.57e+02	617	5	25.0	724	1	TAP1_MOUSE	ANTIGEN PEPTIDE TRANSP	2.57e+02
545	5	25.0	596	1	SDP_EIMBO	PHENYLLANILY-L-TRNA SYN	2.57e+02	618	5	25.0	725	1	TAP1_RAT	ANTIGEN PEPTIDE TRANSP	2.57e+02
546	5	25.0	596	1	TR4_RAT	SPOROZOITE DEVELOPMENT	2.57e+02	619	5	25.0	726	1	ATL_CAMPC	ANTIGEN PEPTIDE TRANSP	2.57e+02
547	5	25.0	596	1	TR4_MOUSE	ORPHAN RECEPTOR TR4	2.57e+02	620	5	25.0	727	1	CATA_SALTY	A-TYPE INCLUSION PROTE	2.57e+02
548	5	25.0	596	1	TR4_HUMAN	ORPHAN RECEPTOR TR4	2.57e+02	621	5	25.0	727	1	CATA_SALTY	CATALASE HPI (EC 1.11.	2.57e+02
549	5	25.0	598	1	YAU6_SCHPO	ORPHAN RECEPTOR TR4	2.57e+02	622	5	25.0	728	1	YJ89_YEAST	HYPOTHETICAL 95.0 KD P	2.57e+02
550	5	25.0	598	1	BGAL_XANMN	HYPOTHETICAL 60.7 KD P	2.57e+02	623	5	25.0	730	1	FIXI_BRAJA	HYPOTHETICAL 85.0 KD P	2.57e+02
551	5	25.0	602	1	DCPI_ORYSA	BETA-GALACTOSIDASE PRE	2.57e+02	633	5	25.0	735	1	DUR3_YEAST	NITROGEN FIXATION PROT	2.57e+02
552	5	25.0	603	1	TBFG_EPTST	PYRUVATE DECARBOXYLASE	2.57e+02	624	5	25.0	738	1	HLV1_VIBCH	UREA ACTIVE TRANSPORT	2.57e+02
553	5	25.0	603	1	HIPI_YEAST	THREAD BIOPOLYMER FILA	2.57e+02	625	5	25.0	741	1	HLV2_VIBCH	HEMOLYSIN PRECURSOR.	2.57e+02
554	5	25.0	604	1	CIK9_MOUSE	HISTIDINE PERMEASE.	2.57e+02	626	5	25.0	741	1	IDH2_VIBAI	HAEMOLYSIN PRECURSOR.	2.57e+02
555	5	25.0	605	1	ILVD_ECOLI	VOLTAGE-GATED POTASSIU	2.57e+02	627	5	25.0	744	1	LEU2_HYBL	ISOCITRATE DEHYDROGENA	2.57e+02
556	5	25.0	607	1	RIB1_RAT	DIHYDROXY-ACID DEHYDRA	2.57e+02	628	5	25.0	745	1	CUL2_HUMAN	3-ISOPROPYLMALATE DEHY	2.57e+02
557	5	25.0	607	1	RIB1_HUMAN	DOLICHYL-DIPHOSPHOOLIG	2.57e+02	629	5	25.0	747	1	YN54_CAEEL	CULLIN HOMOLOG 2 (CUL-	2.57e+02
558	5	25.0	609	1	FES_FSVGA	DOLICHYL-DIPHOSPHOOLIG	2.57e+02	630	5	25.0	747	1	EZH1_HUMAN	HYPOTHETICAL 80.8 KD P	2.57e+02
559	5	25.0	610	1	SPPA_NY3	TYROSINE-PROTEIN KINAS	2.57e+02	631	5	25.0	747	1	EZH1_MOUSE	ENHANCER OF ZESTE HOMO	2.57e+02
560	5	25.0	610	1	DCPI_MAIZE	PROTEASE IV HOMOLOG (E	2.57e+02	632	5	25.0	748	1	TAP1_HUMAN	ENHANCER OF ZESTE HOMO	2.57e+02
561	5	25.0	611	1	XYNA_PSEFL	PYRUVATE DECARBOXYLASE	2.57e+02	633	5	25.0	752	1	CUL1_HUMAN	ANTIGEN PEPTIDE TRANSP	2.57e+02
562	5	25.0	612	1	ILVD_HAEN	ENDO-1,4-BETA-XYLANASE	2.57e+02	634	5	25.0	758	1	NSFH_CAEEL	CULLIN HOMOLOG 1 (CUL-	2.57e+02
563	5	25.0	613	1	CIK5_HUMAN	DIHYDROXY-ACID DEHYDRA	2.57e+02	635	5	25.0	768	1	YJ8D_ECOLI	VESICULAR-FUSION PROTE	2.57e+02
564	5	25.0	615	1	NIFA_AZOCA	VOLTAGE-GATED POTASSIU	2.57e+02	636	5	25.0	768	1	TREG_ASPAW	HYPOTHETICAL 86.7 KD P	2.57e+02
565	5	25.0	615	1	SECD_ECOLI	NIF-SPECIFIC REGULATOR	2.57e+02	637	5	25.0	773	1	LI19_DROME	ANTHRANILATE SYNTHASE	2.57e+02
566	5	25.0	615	1	SPPA_HAEN	PROTEIN-EXPORT MEMBRAN	2.57e+02	638	5	25.0	774	1	SCA_DROME	LIN-19 HOMOLOG PROTEIN	2.57e+02
567	5	25.0	618	1	LEUL_YEAST	PROTEASE IV (EC 3.4.-	2.57e+02	639	5	25.0	775	1	ICP0_HSV11	SCABROUS PROTEIN PRECU	2.57e+02
568	5	25.0	619	1	HSCA_HAEN	2-ISOPROPYLMALATE SYN	2.57e+02	640	5	25.0	780	1	CUL3_CAEEL	TRANS-ACTING TRANSCRIP	2.57e+02
569	5	25.0	620	1	EXON_HSV2	CHAPERONE PROTEIN HSCA	2.57e+02	641	5	25.0	784	1	GCF_HUMAN	CUL-3 PROTEIN.	2.57e+02
570	5	25.0	621	1	VP40_HSVBC	ALKALINE EXONUCLEASE (	2.57e+02	642	5	25.0	785	1	CUL3_SCHPO	CUL-RICH SEQUENCE CNA-B	2.57e+02
571	5	25.0	624	1	VP87_NSVOP	CAPSID PROTEIN P40 [CO	2.57e+02	643	5	25.0	785	1	DMSA_ECOLI	CULLIN HOMOLOG 3 (CUL-	2.57e+02
572	5	25.0	627	1	SPD2_NEPCL	CAPSID PROTEIN P87.	2.57e+02	644	5	25.0	787	1	RIR1_HSVBC	ANEROBIC DIMETHYL SUL	2.57e+02
573	5	25.0	637	1	Y561_HAEN	SPIRODIN 2 (DRAGLINE S	2.57e+02	645	5	25.0	789	1	K25_CAEEL	RIBONUCLEOSIDE-DIPHOS	2.57e+02
574	5	25.0	637	1	Y561_HAEN	HYPOTHETICAL PROTEIN H	2.57e+02	646	5	25.0	789	1	K6PF_HAECO	HYPOTHETICAL 87.8 KD P	2.57e+02
575	5	25.0	633	1	DPX2_BACTK	70 KD CRYSTAL PROTEIN	2.57e+02	647	5	25.0	790	1	TNP3_HUMAN	6-PHOSPHOFRUCTOKINASE	2.57e+02
576	5	25.0	637	1	ODP2_AZOVI	DIPHOSOLIPOMIDE ACETY	2.57e+02	648	5	25.0	803	1	CUL4_CAEEL	TUMOR NECROSIS FACTOR,	2.57e+02
577	5	25.0	640	1	PPCM_CHICK	PHOSPHOENOLPYRUVATE (	2.57e+02	649	5	25.0	803	1	GYRB_SALTY	CUL-4 PROTEIN.	2.57e+02
578	5	25.0	641	1	TTK8_DROME	PHOSPHOENOLPYRUVATE CA	2.57e+02	650	5	25.0	803	1	GYRB_ECOLI	DNA GYRASE SUBUNIT B (	2.57e+02
579	5	25.0	642	1	DP3X_SALTY	TRAMTRACK PROTEIN, BET	2.57e+02	651	5	25.0	804	1	YBBP_ECOLI	DNA GYRASE SUBUNIT B (	2.57e+02
580	5	25.0	643	1	PGT_RAT	DNA POLYMERASE III SUB	2.57e+02	652	5	25.0	806	1	FES_FELCA	DNA GYRASE SUBUNIT B (	2.57e+02
581	5	25.0	643	1	PGT_HUMAN	PROSTAGLANDIN TRANSPOR	2.57e+02	653	5	25.0	820	1	FES_FELCA	PROTO-ONCOGENE TYROSIN	2.57e+02
582	5	25.0	646	1	UVRC_MYCTU	EXCINUCLEASE ABC SUBUN	2.57e+02	654	5	25.0	821	1	GYRA_BACSU	DNA GYRASE SUBUNIT A (	2.57e+02
583	5	25.0	647	1	GAG_SFVI	GAG POLYPROTEIN (CORE	2.57e+02	655	5	25.0	822	1	FES_HUMAN	PROTO-ONCOGENE TYROSIN	2.57e+02
584	5	25.0	648	1	RAD4_SCHPO	RAD4 PROTEIN (CUTS PRO	2.57e+02	656	5	25.0	826	1	LOX2_MTXXA	ATP-DEPENDENT PROTEASE	2.57e+02
585	5	25.0	652	1	DREB_CHICK	DREBRINS A, E1 AND E2.	2.57e+02	657	5	25.0	827	1	SOX6_MOUSE	TRANSCRIPTION FACTOR S	2.57e+02
586	5	25.0	652	1	MX1_RAT	INTERFERON-INDUCED GTP	2.57e+02	658	5	25.0	831	1	DP01_THEFL	DNA POLYMERASE I, THER	2.57e+02
587	5	25.0	653	1	APL_ACHLY	PROTEASE I PRECURSOR (	2.57e+02	659	5	25.0	837	1	NCM2_HUMAN	NEURAL CELL ADHESION M	2.57e+02
588	5	25.0	654	1	BGL_ERWCH	PERILASMIC BETA-GLUCO	2.57e+02	660	5	25.0	841	1	PSPI_YEAST	PSPI PROTEIN (GIN5 PRO	2.57e+02
589	5	25.0	657	1	MYB_DROME	MYB PROTEIN.	2.57e+02	661	5	25.0	850	1	DP01_ANATH	DNA POLYMERASE I (EC 2	2.57e+02
590	5	25.0	659	1	VST2_HEVPA	STRUCTURAL PROTEIN 2 P	2.57e+02	662	5	25.0	852	1	CLPB_CORGL	CLPB PROTEIN.	2.57e+02
591	5	25.0	660	1	AMY_BACSU	STRUCTURAL PROTEIN 2 P	2.57e+02	663	5	25.0	859	1	PPS3_BACSU	PEPTIDE SYNTHETASE 3 (	2.57e+02
592	5	25.0	660	1	AMY_HUMAN	ALPHA-AMYLASE PRECURSO	2.57e+02	664	5	25.0	865	1	CPN_DROME	CALPHOTIN.	2.57e+02
593	5	25.0	661	1	MX1_HUMAN	INTERFERON-REGULATED R	2.57e+02	665	5	25.0	869	1	Y189_MYCLE	HYPOTHETICAL 95.7 KD P	2.57e+02
594	5	25.0	662	1	TRB2_RHISN	PROBABLE CONJUGAL TRAN	2.57e+02	666	5	25.0	881	1	HELL_HSV2H	PROBABLE HELICASE.	2.57e+02
595	5	25.0	666	1	GLGE_BACCL	1,4-ALPHA-GLUCAN BRANC	2.57e+02	667	5	25.0	882	1	HELL_HSV11	PROBABLE HELICASE.	2.57e+02
596	5	25.0	667	1	YVAE_BACSU	HYPOTHETICAL 73.9 KD P	2.57e+02	668	5	25.0	886	1	CHS_SAPMO	CHITIN SYNTHASE (EC 2.	2.57e+02
597	5	25.0	671	1	ICP0_FCVF9	COAT PROTEIN (CAPSID P	2.57e+02	669	5	25.0	887	1	OBP_HSVB	REPLICATION ORIGIN BIN	2.57e+02
598	5	25.0	676	1	ICP0_HSVB	TRANS-ACTING TRANSCRIP	2.57e+02	670	5	25.0	890	1	IF2_ECOLI	TRANSLATION INITIATION	2.57e+02
599	5	25.0	676	1	ICP0_HSVB	TRANS-ACTING TRANSCRIP	2.57e+02	671	5	25.0	894	1	WPR4_BACSU	CELL WALL-ASSOCIATED P	2.57e+02
600	5	25.0	686	1	CHEA_HRSH	CHROMATIN PROTEIN CHE	2.57e+02	672	5	25.0	905	1	HEX_ADECR	HEXON PROTEIN (LATE PR	2.57e+02
601	5	25.0	690	1	RHO_MICLU	CHROMATIN PROTEIN CHE	2.57e+02	673	5	25.0	905	1	HEX_ADECR	HEXON PROTEIN (LATE PR	2.57e+02
602	5	25.0	691	1	COMA_NEIGO	COMPTENCE PROTEIN COM	2.57e+02	674	5	25.0	906	1	KBF2_CHICK	NUCLEAR FACTOR NF-KAPP	2.57e+02
603	5	25.0	697	1	TRP_SCHPO	TRYPTOPHAN SYNTHASE (E	2.57e+02	675	5	25.0	908	1	SRCA_RABIT	SARCALUMENIN PRECURSOR	2.57e+02
604	5	25.0	699	1	PTGA_BACSU	PTS SYSTEM, GLUCOSE-SP	2.57e+02	676	5	25.0	920	1	PM1_NEUCR	PLASMA MEMBRANE APPASE	2.57e+02
605	5	25.0	699	1	VGUG_HSV2H	GLYCOPROTEIN G.	2.57e+02	677	5	25.0	921	1	T2D3_DROME	TRANSCRIPTION INITIATI	2.57e+02
606	5	25.0	700	1	NUOL_RHOCA	NADH DEHYDROGENASE I C	2.57e+02	678	5	25.0	923	1	PRGR_MOUSE	PROCESSTRONE RECEPTO	2.57e+02
607	5	25.0	700	1	PALY_WHEAT	PHENYLLANINE AMMONIA-	2.57e+02	680	5	25.0	935	1	OD01_HAEN	2-OXOGLUTARATE DEHYDR	2.57e+02
											936	1	CAPP_RHOOB	PHOSPHOENOLPYRUVATE CA	2.57e+02

681	5	25.0	936	1	FHL1_YEAST	PRE-RRNA PROCESSING PR	2.57e+02	754	5	25.0	1646	1	RRPO_CGMVS	PUTATIVE RNA-DIRECTED	2.57e+02
682	5	25.0	937	1	PAC4_RAT	SUBTILISIN-LIKE PROTEA	2.57e+02	755	5	25.0	1677	1	VIT_ACITR	VITELLOGENIN PRECURSOR	2.57e+02
683	5	25.0	949	1	TRH_DROME	EXCINUCLEASE PROTEIN.	2.57e+02	756	5	25.0	1690	1	RPOC_THEMA	DNA-DIRECTED RNA POLYM	2.57e+02
684	5	25.0	950	1	UVPA_NEIGO	EXCINUCLEASE ABC SUBUN	2.57e+02	757	5	25.0	1752	1	RPB1_SCHPO	DNA-DIRECTED RNA POLYM	2.57e+02
685	5	25.0	960	1	YAO2_SCHPO	HYPOTHETICAL 108.7 KD	2.57e+02	758	5	25.0	1788	1	POLN_SOUV3	NON-STRUCTURAL POLYPRO	2.57e+02
686	5	25.0	962	1	GUNA_PSEFL	ENDOGLUCANASE A PRECUR	2.57e+02	759	5	25.0	1885	1	FAS2_CANAL	FATTY ACID SYNTHASE, S	2.57e+02
687	5	25.0	966	1	M172_HUMAN	MEMBRANE COMPONENT, CH	2.57e+02	760	5	25.0	2017	1	MYSN_DROME	MYOSIN HEAVY CHAIN, NO	2.57e+02
688	5	25.0	967	1	SOXA_CORSP	SARCOSINE OXIDASE ALPH	2.57e+02	761	5	25.0	2214	1	PYR1_YEAST	URA2 PROTEIN [CONTAINS	2.57e+02
689	5	25.0	969	1	PAC4_HUMAN	SUBTILISIN-LIKE PROTEA	2.57e+02	762	5	25.0	2363	1	SPCO_MOUSE	SPECTRIN BETA CHAIN, B	2.57e+02
690	5	25.0	971	1	RPAL_HALHA	DNA-DIRECTED RNA POLYM	2.57e+02	763	5	25.0	2414	1	P300_HUMAN	E1A-ASSOCIATED PROTEIN	2.57e+02
691	5	25.0	982	1	CAPP_ANASP	PHOSPHOENOLPYRUVATE CA	2.57e+02	764	5	25.0	2431	1	POLN_SFV	NONSTRUCTURAL POLYPROT	2.57e+02
692	5	25.0	984	1	DPOL_NPVAC	DNA POLYMERASE (EC 2.7	2.57e+02	765	5	25.0	2479	1	POLN_RRVN	NONSTRUCTURAL POLYPROT	2.57e+02
693	5	25.0	985	1	ENV_SFV1	ENV POLYPROTEIN [COAT	2.57e+02	766	5	25.0	2560	1	PPS2_BACSU	PEPTIDE SYNTHETASE 2.	2.57e+02
694	5	25.0	994	1	ENV_FOAMV	ENV POLYPROTEIN (COAT	2.57e+02	767	5	25.0	2594	1	7LES_DROVI	SEVENTHLESS PROTEIN (EC	2.57e+02
695	5	25.0	986	1	DPOL_NPVBM	DNA POLYMERASE (EC 2.7	2.57e+02	768	5	25.0	2701	1	IP3S_HUMAN	INOSITOL 1,4,5-TRISPHO	2.57e+02
696	5	25.0	986	1	GNF1_DROME	GERMLINE TRANSCRIPTION	2.57e+02	769	5	25.0	2701	1	IP3S_RAT	INOSITOL 1,4,5-TRISPHO	2.57e+02
697	5	25.0	991	1	DHP1_SCHPO	DHP1 PROTEIN.	2.57e+02	770	5	25.0	2749	1	IP3R_MOUSE	INOSITOL 1,4,5-TRISPHO	2.57e+02
698	5	25.0	997	1	DPOL_TREPA	DNA POLYMERASE I (EC 2	2.57e+02	771	5	25.0	2749	1	IP3R_RAT	INOSITOL 1,4,5-TRISPHO	2.57e+02
699	5	25.0	1001	1	IF2_SINY3	TRANSLATION INITIATION	2.57e+02	772	5	25.0	2833	1	IP3R_DROME	INOSITOL 1,4,5-TRISPHO	2.57e+02
700	5	25.0	1004	1	GCSP_CHICK	GLYCINE DEHYDROGENASE	2.57e+02	773	5	25.0	3005	1	POLG_TVIV	GENOME POLYPROTEIN [CO	2.57e+02
701	5	25.0	1014	1	HEX2_YEAST	HEX2 PROTEIN (SRN1 PRO	2.57e+02	774	5	25.0	3027	1	POLG_PYFV1	GENOME POLYPROTEIN [CO	2.57e+02
702	5	25.0	1017	1	DPOL_ADEB2	DNA POLYMERASE (EC 2.7	2.57e+02	775	5	25.0	3061	1	POLG_PVYHU	GENOME POLYPROTEIN [CO	2.57e+02
703	5	25.0	1030	1	SPAB_BACSU	SUBTILIN BIOSYNTHESIS	2.57e+02	776	5	25.0	3063	1	POLG_PVYN	GENOME POLYPROTEIN [CO	2.57e+02
704	5	25.0	1034	1	BGAL_KLEPN	BETA-GALACTOSIDASE [EC	2.57e+02	777	5	25.0	3137	1	CA36_CHICK	COLLAGEN ALPHA 3(VI) C	2.57e+02
705	5	25.0	1036	1	YG35_YEAST	HYPOTHETICAL 117.0 KD	2.57e+02	778	5	25.0	3224	1	N358_HUMAN	NUCLEAR PORE COMPLEX P	2.57e+02
706	5	25.0	1043	1	DSG1_BOVIN	DESMOGLEIN 1 PRECURSOR	2.57e+02	779	5	25.0	3567	1	ERY2_SACER	ERYTHRONOLIDE SYNTHASE	2.57e+02
707	5	25.0	1044	1	GPFA_BOVIN	GTPASE-ACTIVATING PROT	2.57e+02	780	5	25.0	3649	1	ACVS_NOCLA	DELTA-(L-ALPHA-AMINOA	2.57e+02
708	5	25.0	1046	1	RPOC_WEIHE	DNA-DIRECTED RNA POLYM	2.57e+02	781	5	25.0	3759	1	TRFX_DROME	TRITHORAX PROTEIN.	2.57e+02
709	5	25.0	1048	1	CPXB_BACME	CYTOCHROME P450(BM-3)	2.57e+02	782	5	25.0	3770	1	ACVS_EMENI	DELTA-(L-ALPHA-AMINOA	2.57e+02
710	5	25.0	1048	1	P100_HCMVA	LARGE STRUCTURAL PHOSP	2.57e+02	783	5	25.0	3866	1	HRX_MOUSE	ZINC FINGER PROTEIN HR	2.57e+02
711	5	25.0	1049	1	DSG1_HUMAN	DESMOGLEIN 1 PRECURSOR	2.57e+02	784	5	25.0	3969	1	HRX_HUMAN	ZINC FINGER PROTEIN HR	2.57e+02
712	5	25.0	1053	1	CAPP_ANANI	PHOSPHOENOLPYRUVATE CA	2.57e+02	785	5	25.0	4488	1	RRPA_CVMJH	RNA-DIRECTED RNA POLYM	2.57e+02
713	5	25.0	1054	1	RPOC_WEIPA	DNA-DIRECTED RNA POLYM	2.57e+02	786	5	25.0	4655	1	LRP2_HUMAN	LOW-DENSITY LIPOPROTEI	2.57e+02
714	5	25.0	1057	1	SPS1_CITUN	SUCROSE-PHOSPHATE SYNT	2.57e+02	787	4	20.0	62	1	COXS_MOUSE	CYTOCHROME C OXIDASE A	4.16e+03
715	5	25.0	1059	1	SPS_VICFA	SUCROSE-PHOSPHATE SYNT	2.57e+02	788	4	20.0	66	1	LHA2_RHOPA	LIGHT-HARVESTING PROTE	4.16e+03
716	5	25.0	1061	1	HIRA_DROME	HIRA PROTEIN HOMOLOG.	2.57e+02	789	4	20.0	70	1	ALPA_ECOLI	PHOPHAGE CP4-57 REGULA	4.16e+03
717	5	25.0	1063	1	PDRL_YEAST	PLEOTROPIC DRUG RESIS	2.57e+02	790	4	20.0	85	1	PYS1_FREDI	PHYCOBILISOME 9.7 KD L	4.16e+03
718	5	25.0	1081	1	KOKO_YEAST	PROBABLE SERINE/THREON	2.57e+02	791	4	20.0	93	1	IAPP_RAT	ISLET AMYLOID POLYPEPT	4.16e+03
719	5	25.0	1088	1	PRH_PETCR	PATHOGENESIS-RELATED H	2.57e+02	792	4	20.0	120	1	HV3U_HUMAN	IG HEAVY CHAIN V-III R	4.16e+03
720	5	25.0	1091	1	NCAL_CHICK	NEURAL CELL ADHESION M	2.57e+02	793	4	20.0	129	1	ACP2_HORVU	ACP2 CARRIER PROTEIN I	4.16e+03
721	5	25.0	1122	1	ADP1_MYCGA	ADHESIN P1 PRECURSOR (	2.57e+02	794	4	20.0	133	1	AO6_YEAST	PROTEIN AO.6 PRECURSOR	4.16e+03
722	5	25.0	1133	1	EGF_RAT	EPIDERMAL GROWTH FACTO	2.57e+02	795	4	20.0	135	1	IL4_SHEEP	INTERLEUKIN-4 PRECURSO	4.16e+03
723	5	25.0	1146	1	YHC3_YEAST	HYPOTHETICAL 130.0 KD	2.57e+02	796	4	20.0	137	1	KDGL_RHIME	DIACYLGLYCEROL KINASE	4.16e+03
724	5	25.0	1146	1	KHMA_DICDI	MYOSIN HEAVY CHAIN KIN	2.57e+02	797	4	20.0	144	1	ATP2_STRLI	ATP SYNTHASE PROTEIN I	4.16e+03
725	5	25.0	1148	1	POLN_RVVT	NONSTRUCTURAL POLYPROT	2.57e+02	798	4	20.0	151	1	GLB_CERHR	GLOBIN (MYOGLOBIN).	4.16e+03
726	5	25.0	1168	1	MYSC_ACACA	MYOSIN IC HEAVY CHAIN.	2.57e+02	799	4	20.0	156	1	TEX1_HUMAN	RADIATION-INDUCIBLE IM	4.16e+03
727	5	25.0	1173	1	NIFJ_ENTAG	PIRUVATE-FLAVODOXIN OX	2.57e+02	800	4	20.0	156	1	13MS_BRANA	MICROSPORE-SPECIFIC PR	4.16e+03
728	5	25.0	1178	1	RPOB_TREPA	DNA-DIRECTED RNA POLYM	2.57e+02	801	4	20.0	159	1	DHSD_HUMAN	SUCCINATE DEHYDROGENAS	4.16e+03
729	5	25.0	1181	1	ITAZ2_HUMAN	PLATELET MEMBRANE GLYC	2.57e+02	802	4	20.0	168	1	ID1_MOUSE	DNA-BINDING PROTEIN IN	4.16e+03
730	5	25.0	1192	1	XPG_MOUSE	DNA-REPAIR PROTEIN COM	2.57e+02	803	4	20.0	168	1	TAO3_WHEAT	ALPHA-AMYLASE/TRYPsin	4.16e+03
731	5	25.0	1194	1	BCHH_RHOCA	MAGNESIUM-PROTOPORPHYR	2.57e+02	804	4	20.0	173	1	CRAA_ZALCA	ALPHA CRYSTALLIN A CHA	4.16e+03
732	5	25.0	1196	1	XPG_XENLA	DNA-REPAIR PROTEIN COM	2.57e+02	805	4	20.0	173	1	CRAA_URSUR	ALPHA CRYSTALLIN A CHA	4.16e+03
733	5	25.0	1216	1	AGEP_RAT	APICAL ENDOSOMAL GLICO	2.57e+02	806	4	20.0	173	1	CRAA_RABIT	ALPHA CRYSTALLIN A CHA	4.16e+03
734	5	25.0	1223	1	YWR1_CAEEL	PUTATIVE TYROSINE-PROT	2.57e+02	807	4	20.0	173	1	CRAA_HORSE	ALPHA CRYSTALLIN A CHA	4.16e+03
735	5	25.0	1248	1	TOPG_SULAC	REVERSE GYRASE [CONTAI	2.57e+02	808	4	20.0	173	1	CRAA_ALLMI	ALPHA CRYSTALLIN A CHA	4.16e+03
736	5	25.0	1271	1	BCR_HUMAN	BREAKPOINT CLUSTER REG	2.57e+02	809	4	20.0	176	1	ITRY_ACACO	TRYPsin INHIBITOR PREC	4.16e+03
737	5	25.0	1284	1	ATI_COMPX	A-TYPE INCLUSION PROTE	2.57e+02	810	4	20.0	182	1	FTN_DROME	FLIGHTIN (MUSCLE PROTE	4.16e+03
738	5	25.0	1298	1	ICP4_HSV11	TRANS-ACTING TRANSCRIP	2.57e+02	811	4	20.0	185	1	KITH_CPEPV	THYMIDINE KINASE (EC 2	4.16e+03
739	5	25.0	1319	1	DYNA_DROME	150 KD DYNEIN-ASSOCIAT	2.57e+02	812	4	20.0	187	1	R12C_ARATH	50S RIBOSOMAL PROTEIN	4.16e+03
740	5	25.0	1327	1	YML1_MYCTU	HYPOTHETICAL 145.8 KD	2.57e+02	813	4	20.0	196	1	CRA2_SPAEH	ALPHA CRYSTALLIN A CHA	4.16e+03
741	5	25.0	1356	1	Y195_HUMAN	HYPOTHETICAL REPAIR PR	2.57e+02	814	4	20.0	202	1	COAT_BTWVG	ALPHA CRYSTALLIN A CHA	4.16e+03
742	5	25.0	1358	1	MSH6_MOUSE	DNA MISMATCH REPAIR K	2.57e+02	815	4	20.0	211	1	GT28_SCHBO	GLUTATHIONE S-TRANSFER	4.16e+03
743	5	25.0	1365	1	SUZ2_DROME	SUPPRESSOR 2 OF ZESTE	2.57e+02	816	4	20.0	212	1	IL6_MACFA	INTERLEUKIN-6 PRECURSO	4.16e+03
744	5	25.0	1408	1	GEA1_YEAST	ARF GUANINE-NUCLEOTIDE	2.57e+02	817	4	20.0	214	1	KAD_SALTY	ADENYLATE KINASE (EC 2	4.16e+03
745	5	25.0	1442	1	UY9F_YEAST	HYPOTHETICAL 161.2 KD	2.57e+02	818	4	20.0	216	1	LEF2_NPVLD	LATE EXPRESSION FACTOR	4.16e+03
746	5	25.0	1487	1	ICP4_HSVK	TRANS-ACTING TRANSCRIP	2.57e+02	819	4	20.0	218	1	COAT_CMWDK	COAT PROTEIN (CAPSID P	4.16e+03
747	5	25.0	1487	1	ICP4_HSVK	TRANS-ACTING TRANSCRIP	2.57e+02	820	4	20.0	218	1	COAT_CMWDK	COAT PROTEIN (CAPSID P	4.16e+03
748	5	25.0	1489	1	YGP0_YEAST	HYPOTHETICAL 171.5 KD	2.57e+02	821	4	20.0	222	1	ARTM_ECOLI	ARGININE TRANSPORT SYS	4.16e+03
749	5	25.0	1545	1	MRP2_HUMAN	CANALICULAR MULTISPECI	2.57e+02	822	4	20.0	224	1	KCY_BACSU	CYTIDYLATE KINASE (EC	4.16e+03
750	5	25.0	1559	1	STCJ_EMENI	PUTATIVE STERIGMATOCYS	2.57e+02	823	4	20.0	225	1	DEOC_SINY3	DEOXYRIBOSE-PHOSPHATE	4.16e+03
751	5	25.0	1588	1	AR01_YEAST	PENTAFUNCTIONAL AROM P	2.57e+02	824	4	20.0	239	1	FLIA_ECOLI	RNA POLYMERASE SIGMA F	4.16e+03
752	5	25.0	1596	1	GLI3_HUMAN	ZINC FINGER PROTEIN GL	2.57e+02	825	4	20.0	243	1	MUKE_HAEIN	MUKE PROTEIN HOMOLOG.	4.16e+03
753	5	25.0	1603	1	AR01_EMENI	PENTAFUNCTIONAL AROM P	2.57e+02	826	4	20.0	246	1	IF6_CAEEL	EUKARYOTIC TRANSLATION	4.16e+03



827	4	20.0	1	ETRA_SHEPU	ELECTRON TRANSPORT REG	4.16e+03	900	1	ABFB_STRLI	ALPHA-L-ARABINOFURANOS	4.16e+03
828	4	20.0	1	APAH_KLEAE	BIS(5'-NUCLEOSYL)-TEIR	4.16e+03	901	1	RBL_GOSHI	RIBULOSE BISPHOSPHATE	4.16e+03
829	4	20.0	1	143E_DROME	14-3-3 PROTEIN EPSILON	4.16e+03	902	1	LEGU_CANEN	LEGUMIN PRECURSOR (HC	4.16e+03
830	4	20.0	1	COAT_TNVD	COAT PROTEIN	4.16e+03	903	1	HRPB_BURSO	REGULATORY PROTEIN RPB	4.16e+03
831	4	20.0	1	HUPH_HAOCA	HYDROGENASE EXPRESSION	4.16e+03	904	1	DLID2_PSEPU	LIPAMIDE DEHYDROGENAS	4.16e+03
832	4	20.0	1	ACCD_HAEIN	ACETYL-COENZYME A CARB	4.16e+03	905	1	GLGA_AGRTU	GLYCOCEN SYNTHASE (EC	4.16e+03
833	4	20.0	1	BPRC_PSEPA	BIPHENYL-2,3-DIOL 1,2-	4.16e+03	906	1	KNIR_DROVI	ZYGOTIC GAP PROTEIN KN	4.16e+03
834	4	20.0	1	KHSE_CORGL	HOMOSERINE KINASE (EC	4.16e+03	907	1	CON2_BPPF3	MINOR COAT PROTEIN (OR	4.16e+03
835	4	20.0	1	KHSE_BRELA	HOMOSERINE KINASE (EC	4.16e+03	908	1	GTR5_RABIT	GLUCOSE TRANSPORTER TY	4.16e+03
836	4	20.0	1	IF2B_SCHPO	PROBABLE EUKARYOTIC TR	4.16e+03	909	1	LEF9_NPVOP	LATE EXPRESSION FACTOR	4.16e+03
837	4	20.0	1	CDX2_MOUSE	HOMEOBOX PROTEIN CDX-2	4.16e+03	910	1	ACCD_EPIVI	ACETYL-COENZYME A CARB	4.16e+03
838	4	20.0	1	CDX1_KULJA	CYTCHROME C OXIDASE A	4.16e+03	911	1	ACCD_CUSRE	ACETYL-COENZYME A CARB	4.16e+03
839	4	20.0	1	CDX2_MESAU	HOMEOBOX PROTEIN CDX-2	4.16e+03	912	1	CPN1_MESAU	CYTCHROME P450 XIB1 P	4.16e+03
840	4	20.0	1	KDGR_HAEIN	2-DEHYDRO-3-DEOXYGLUCO	4.16e+03	913	1	CPN1_HUMAN	CYTCHROME P450 XIB1 P	4.16e+03
841	4	20.0	1	DHYG_HYPME	GLUCURONATE DEHYDROGENAS	4.16e+03	914	1	CEAB_ECOLI	GMP SYNTHASE [GLUTAMIN	4.16e+03
842	4	20.0	1	IF34_DROME	EUKARYOTIC TRANSLATION	4.16e+03	915	1	CP11_RABIT	CYCLOXIMETHYLGLUTARYL-	4.16e+03
843	4	20.0	1	IF34_SCHPO	EUKARYOTIC TRANSLATION	4.16e+03	916	1	HMC5_HUMAN	HYDROXYMETHYLGLUTARYL-	4.16e+03
844	4	20.0	1	HEM2_STRCO	DELTA-AMINOLEVULINIC A	4.16e+03	917	1	GLG2_WHEAT	GLUCOSE-1-PHOSPHATE AD	4.16e+03
845	4	20.0	1	ASPG_WOLSU	L-ASPARAGINASE (EC 3.5	4.16e+03	918	1	MPPA_RAT	MITOCHONDRIAL PROCESSI	4.16e+03
846	4	20.0	1	GCAA_MOUSE	IG GAMMA-2A CHAIN C RE	4.16e+03	919	1	GUAU_MYCTU	GMP SYNTHASE (GLUTAMIN	4.16e+03
847	4	20.0	1	DMPL_PSESP	PHENOL HYDROXYLASE PI	4.16e+03	920	1	CIK3_RAT	VOLTAGE-GATED POTASSIU	4.16e+03
848	4	20.0	1	AP42_KULJA	5',5''',-P-1,P-4-TETRAP	4.16e+03	921	1	MPPA_HUMAN	MITOCHONDRIAL PROCESSI	4.16e+03
849	4	20.0	1	HMSU_YERPE	HEMIN TRANSPORT PROTEI	4.16e+03	922	1	HEXA_DICDI	BETA-HEXOSAMINIDASE AL	4.16e+03
850	4	20.0	1	AROG_ECOLI	PHOSPHO-2-DEHYDRO-3-DE	4.16e+03	923	1	GNIB_RAT	CALCIUM/CALMODULIN-DSP	4.16e+03
851	4	20.0	1	LEGR_PEA	LEGUMIN K (FRAGMENT)	4.16e+03	924	1	HSF2_HUMAN	HEAT SHOCK FACTOR PROT	4.16e+03
852	4	20.0	1	GLNA_AGABI	GLUTAMINE SYNTHETASE (	4.16e+03	925	1	COXL_BRAJA	CYCLOXIMETHYLGLUTARYL-	4.16e+03
853	4	20.0	1	KARG_SCHAM	ARGININE KINASE (EC 2.4	4.16e+03	926	1	KPYR_MOUSE	PYRUVATE KINASE, ISOZY	4.16e+03
854	4	20.0	1	LEU3_SOLTU	3-ISOPROPYLMALATE DEHY	4.16e+03	927	1	HMD1_HEVBR	3-HYDROXY-3-METHYLGLUT	4.16e+03
855	4	20.0	1	LEU3_THIFE	3-ISOPROPYLMALATE DEHY	4.16e+03	928	1	ILIR_MOUSE	INTERLEUKIN-1 RECEPTO	4.16e+03
856	4	20.0	1	FCGN_RAT	IGG RECEPTOR FCGRN LARG	4.16e+03	929	1	ILV5_SPIOL	KETOL-ACID REDUCTOISOM	4.16e+03
857	4	20.0	1	ALFI_CAEEL	FRUCTOSE-BISPHOSPHATE	4.16e+03	930	1	ALAD_RABIT	ALPHA-ID ADRENERGIC RE	4.16e+03
858	4	20.0	1	CEI1_HORVU	CYSTEINE PROTEINASE EP	4.16e+03	931	1	ILV5_SPIOL	KETOL-ACID REDUCTOISOM	4.16e+03
859	4	20.0	1	CE10_CHICK	CEF-10 PROTEIN PRECURS	4.16e+03	932	1	GNIB_MOUSE	MAJOR CENTROMERE AUTOL	4.16e+03
860	4	20.0	1	ACT1_BOMMO	ACTIN, MUSCLE A1.	4.16e+03	933	1	GPI1_YEAST	N-ACETYLGLUCOSAMINYL-P	4.16e+03
861	4	20.0	1	ACT1_ARTSX	ACTIN, CLONE 205.	4.16e+03	934	1	LN-9 CAEEL	LIN-9 PROTEIN.	4.16e+03
862	4	20.0	1	ADH1_ARATH	ADP-ATP CARRIER PROTEI	4.16e+03	935	1	ACE_DROME	ANGIOTENSIN CONVERTING	4.16e+03
863	4	20.0	1	MYB_AVIMB	GLUTATHIONE-DEPENDENT	4.16e+03	936	1	HS70_CHICK	HEAT SHOCK 70 KD PROTE	4.16e+03
864	4	20.0	1	DHSS_ANACY	SOLUBLE HYDROGENASE 42	4.16e+03	937	1	MYB_MOUSE	MYB PROTO-ONCOGENE PRO	4.16e+03
865	4	20.0	1	ADCB_MYCTU	PROBABLE ACYL-COA DEHY	4.16e+03	938	1	HS74_ANOAL	HEAT SHOCK 70 KD PROTE	4.16e+03
866	4	20.0	1	KC21_CHICK	CASEIN KINASE II, ALPH	4.16e+03	939	1	HS70_ONCTS	HEAT SHOCK 70 KD PROTE	4.16e+03
867	4	20.0	1	GCJM_MOUSE	IG GAMMA-1 CHAIN C RES	4.16e+03	940	1	HS70_MAIZE	HEAT SHOCK 70 KD PROTE	4.16e+03
868	4	20.0	1	AA11_BACSU	ASPARTATE AMINOTRANSF	4.16e+03	941	1	HS70_CHLRE	HEAT SHOCK 70 KD PROTE	4.16e+03
869	4	20.0	1	EFTU_NEIGO	ELONGATION FACTOR TU (	4.16e+03	942	1	HS73_YEAST	HEAT SHOCK 70 KD PROTE	4.16e+03
870	4	20.0	1	ER19_YEAST	DIPHOSPHOMEVALONATE DE	4.16e+03	943	1	HS70_LEIAM	HEAT SHOCK 70 KD PROTE	4.16e+03
871	4	20.0	1	K1CS_HUMAN	KERATIN, TYPE I CYTOSK	4.16e+03	944	1	HS71_TRYCR	MITOCHONDRIAL HEAT SHO	4.16e+03
872	4	20.0	1	GCBM_MOUSE	IG GAMMA-2B CHAIN C RE	4.16e+03	945	1	LAMA_RAT	LAMIN A.	4.16e+03
873	4	20.0	1	CG2B_PATVU	G2/MITOTIC-SPECIFIC CY	4.16e+03	946	1	ILVB_ARATH	HYDROGENASE-4 COMPOEN	4.16e+03
874	4	20.0	1	HXD3_MOUSE	CARBOXYPEPTIDASE A2 PR	4.16e+03	947	1	KFYB_ECOLI	KINESIN-LIKE PROTEIN K	4.16e+03
875	4	20.0	1	CPA2_HUMAN	ACTIN-LIKE PROTEIN 3 (	4.16e+03	948	1	CDGT_BACCI	CYCLOMALTODEXTRIN GLUC	4.16e+03
876	4	20.0	1	ARP3_HUMAN	EARLY GROWTH RESPONSE	4.16e+03	949	1	IDH_CORGL	ISOCITRATE DEHYDROGENA	4.16e+03
877	4	20.0	1	EGR2_XENLA	G2/MITOTIC-SPECIFIC CV	4.16e+03	950	1	FEP4_ECOLI	FERRITINROBACTIN RECE	4.16e+03
878	4	20.0	1	CGA2_MOUSE	CORTICOSTEROID 11-BETA	4.16e+03	951	1	MUTA_MOUSE	METHYLMALONYL-COA MUTA	4.16e+03
879	4	20.0	1	DH12_SHEEP	G2/MITOTIC-SPECIFIC CI	4.16e+03	952	1	COR2_PAVPN	COAT PROTEIN VP1 [CONT	4.16e+03
880	4	20.0	1	CG2B_SPIO	ACYL-COA DEHYDROGENASE	4.16e+03	953	1	MUS2_AQAE	MUTS2 PROTEIN.	4.16e+03
881	4	20.0	1	ACDB_HUMAN	ANTITHROMBIN-III (ATII	4.16e+03	954	1	IF2_ENTFC	TRANSLATION INITIATION	4.16e+03
882	4	20.0	1	ANT3_BOVIN	AAC-RICH MRNA CLONE AA	4.16e+03	955	1	DHG_ECOLI	GLUCOSE DEHYDROGENASE	4.16e+03
883	4	20.0	1	AAC3_DICDI	MYC PROTO-ONCOGENE PRO	4.16e+03	956	1	AFSK_STRCO	SERINE/THREONINE PROTE	4.16e+03
884	4	20.0	1	MYC_SHEEP	GLUTAMINE SYNTHETASE (	4.16e+03	957	1	C1CH_TORNA	CHLORIDE CHANNEL PROTE	4.16e+03
885	4	20.0	1	GLNA_THEME	RIBULOSE BISPHOSPHATE	4.16e+03	958	1	ARM_MUSDO	ARMADILLO SEGMENT POLA	4.16e+03
886	4	20.0	1	EBL1_BYILI	C-ETYS-1 PROTEIN (P54)	4.16e+03	959	1	ERK5_HUMAN	EXTRACELLULAR SIGNAL-R	4.16e+03
887	4	20.0	1	RBL1_HUMAN	PHOSPHATIDATE CYTIDILY	4.16e+03	960	1	DPOL_THETH	DNA POLYMERASE I, THER	4.16e+03
888	4	20.0	1	CD5A_DROME	ORNITHINE DECARBOXYLAS	4.16e+03	961	1	EF2_SCHPO	ELONGATION FACTOR 2 (E	4.16e+03
889	4	20.0	1	DCOR_CRIGR	RXR TYPE HORMONE RECEP	4.16e+03	962	1	IRS1_HCMVA	HYPOTHETICAL PROTEIN H	4.16e+03
890	4	20.0	1	CFL_BOMMO	ADENYLOSUCCINATE LYASE	4.16e+03	963	1	DYN3_RAT	DYNAMIN 3 (DYNAMIN, TE	4.16e+03
891	4	20.0	1	DL03_MOUSE	DIHYDROLIPOAMIDE DEHYD	4.16e+03	964	1	NCAL_HUMAN	NEURAL CELL ADHESION M	4.16e+03
892	4	20.0	1	PUR8_METJA	EARLY GROWTH RESPONSE	4.16e+03	965	1	LDLR_HUMAN	LOW-DENSITY LIPOPROTEI	4.16e+03
893	4	20.0	1	DL03_PSEPU	EARLY GROWTH RESPONSE	4.16e+03	966	1	MYSP_BRUMA	PARMYOSIN.	4.16e+03
894	4	20.0	1	EGR2_MOUSE	GLUTAMINE SYNTHETASE (	4.16e+03	967	1	GLR1_MOUSE	GLUTAMATE RECEPTOR 1 P	4.16e+03
895	4	20.0	1	EGR2_MOUSE	GLUTAMINE SYNTHETASE (	4.16e+03	968	1	PTN4_HUMAN	PROTEIN-TYROSINE PHOSP	4.16e+03
896	4	20.0	1	CISV_CITMA	GLUTAMINE SYNTHETASE (	4.16e+03	969	1	GLR1_MOUSE	GLUTAMATE RECEPTOR 1 P	4.16e+03
897	4	20.0	1	GLNA_SNV3	MITOCHONDRIAL TRIFUNCT	4.16e+03	970	1	PTN4_HUMAN	PROTEIN-TYROSINE PHOSP	4.16e+03
898	4	20.0	1	ECHB_BOVIN	ALPHA-L-ARABINOFURANOS	4.16e+03	971	1	PTN4_HUMAN	PROTEIN-TYROSINE PHOSP	4.16e+03
899	4	20.0	1	ABFE_STRCO	ALPHA-L-ARABINOFURANOS	4.16e+03	972	1	PTN4_HUMAN	PROTEIN-TYROSINE PHOSP	4.16e+03

```
7973 4 20.0 954 1 DRP2_HUMAN DYSTROPHIN-RELATED PRO 4.16e+03
7974 4 20.0 962 1 COPTER BETA SUBUNIT 4.16e+03
7975 4 20.0 976 1 KPS_MOUSE MACROPHAGE COLONY STIM 4.16e+03
7976 4 20.0 1017 1 GSP_SCHPO PUTATIVE GLYCINE DEHYD 4.16e+03
7977 4 20.0 1031 1 BGA2_ECOLI EVOLVED BETA-GALACTOSI 4.16e+03
7978 4 20.0 1034 1 ATHA_PIG POTASSIUM-TRANSPORTING 4.16e+03
7979 4 20.0 1035 1 ATHA_RABIT POTASSIUM-TRANSPORTING 4.16e+03
7980 4 20.0 1057 1 GSP_PEA GLYCINE DEHYDROGENASE 4.16e+03
7981 4 20.0 1072 1 ITA6_CHICK INTEGRIN ALPHA-6 PRECU 4.16e+03
7982 4 20.0 1115 1 NCAL_MOUSE NEURAL CELL ADHESION M 4.16e+03
7983 4 20.0 1178 1 CRW_BACRI 133 KD CRYSTAL PROTEIN 4.16e+03
7984 4 20.0 1201 1 COPA_YEAST COATOMER ALPHA SUBUNIT 4.16e+03
7985 4 20.0 1240 1 B2A2_HUMAN ANION EXCHANGE PROTEIN 4.16e+03
7986 4 20.0 1242 1 IRS1_HUMAN INSULIN RECEPTOR SUBST 4.16e+03
7987 4 20.0 1276 1 MDR3_MOUSE MULTIDRUG RESISTANCE P 4.16e+03
7988 4 20.0 1302 1 MDR5_DROME MULTIDRUG RESISTANCE P 4.16e+03
7989 4 20.0 1312 1 ACE_MOUSE ANGIOTENSIN-CONVERTING 4.16e+03
7990 4 20.0 1394 1 E75B_DROME ECDYSONE-INDUCIBLE PRO 4.16e+03
7991 4 20.0 1444 1 CP5A_BOVIN CLEAVAGE AND POLYADENY 4.16e+03
7992 4 20.0 1468 1 N153_RAT NUCLEAR PORE COMPLEX P 4.16e+03
7993 4 20.0 1502 1 N170_YEAST NUCLEOPORIN NUP170 (NU 4.16e+03
7994 4 20.0 1705 1 CYAA_BORBR CALMODULIN-SENSITIVE A 4.16e+03
7995 4 20.0 2022 1 ANTI_ONCVO MAJOR ANTIGEN. 4.16e+03
7996 4 20.0 2483 1 COA2_HUMAN ACETYL-COA CARBOXYLASE 4.16e+03
7997 4 20.0 4466 1 DYHC_TRIGR DYNEIN BETA CHAIN, CIL 4.16e+03
7998 4 20.0 4563 1 APB_HUMAN APOLIPOPROTEIN B-100 P 4.16e+03
7999 4 20.0 4644 1 DYHC_RAT DYNEIN HEAVY CHAIN, CY 4.16e+03
1000 4 20.0 5179 1 MUC2_HUMAN MUCIN 2 PRECURSOR (INT 4.16e+03

ALIGNMENTS

RESULT 1
ID PUTA_SALTY STANDARD; PRT; 1320 AA.
AC P10503;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROLINE DEHYDROGENASE (EC 1.5.99.8) (PROLINE OXIDASE) / DELTA-1-
DE PYRROLINE-5-CARBOXYLATE DEHYDROGENASE (EC 1.5.1.12) (P5C
DE DEHYDROGENASE).
GN PUTA.
OS SALMONELLA TYPHIMURIUM.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC SALMONELLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93241961.
RA ALLEN S.W., SENTI-WILLIS A.E., MALOY S.R.;
RT "DNA sequence of the puta gene from Salmonella typhimurium: a
RT bifunctional membrane-associated dehydrogenase that binds DNA.";
RL NUCLEIC ACIDS RES. 21:1676-1676(1993).
RN [2]
RP REVISTONS.
RA MALOY S.R.;
RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE; 89127131.
RA HAHN D.R., MYERS R.S., KENT C.R., MALOY S.R.;
RT "Regulation of proline utilization in Salmonella typhimurium:
RT molecular characterization of the put operon, and DNA sequence of the
RT put control region.";
RL MOL. GEN. GENET. 213:125-133(1988).
CC -!- FUNCTION: OXIDIZES PROLINE TO GLUTAMATE FOR USE AS A CARBON AND
CC NITROGEN SOURCE AND ALSO FUNCTION AS A TRANSCRIPTIONAL REPRESSOR
CC OF THE PUT OPERON.
CC -!- CATALYTIC ACTIVITY: L-PROLINE + ACCEPTOR + H(2)O = (S)-1-
CC PYRROLINE-5-CARBOXYLATE + REDUCED ACCEPTOR.
CC -!- CATALYTIC ACTIVITY: 1-PYRROLINE-5-CARBOXYLATE + NAD(+) + H(2)O =
CC L-GLUTAMATE + NADH.
CC -!- COFACTOR: FAD FLAVOPROTEIN.
CC -!- PATHWAY: PROLINE UTILIZATION.
```

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL; Y07627; E284031; -.  
DR EMBL; Y07628; E284031; JOINED.  
DR EMBL; Y07629; E284031; JOINED.  
DR EMBL; Y07630; E284031; JOINED.  
DR EMBL; Y07626; E274594; -.  
DR EMBL; Y07812; E283229; -.  
DR MGD; MGI:892970; MAL.  
KW TRANSMEMBRANE; LIPOPROTEIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16595 MW; 72323D02 CRC32;  
  
Query Match 35.0%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.97e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 3 PAAASGG 9  
QY 14 PAAASGG 20  
  
RESULT 3  
ID MAL\_RAT STANDARD; PRT; 153 AA.  
AC Q64349;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN (17 KD MYELIN VESICULAR  
DE PROTEIN) (MVP17) (NS 3).  
GN MAL.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MORINAE; RAITUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;  
RX MEDLINE; 96129665.  
RA KIM T., FIEDLER K., MADISON D.L., KRUEGER W.H., PEIFFER S.E.;  
RT "Cloning and characterization of MVP17: a developmentally regulated  
RT myelin protein in oligodendrocytes.";  
RL J. NEUROSCI. RES. 42:413-422(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LEWIS; TISSUE-SPINAL CORD;  
RX MEDLINE; 95370938.  
RA SCHAREN-WIEMERS N., VALENZUELA D.M., FRANK M., SCHWAB M.E.;  
RT "Characterization of a rat gene, rMAL, encoding a protein with four  
RT hydrophobic domains in central and peripheral myelin.";  
RL J. NEUROSCI. 15:5753-5764(1995).  
CC -!- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING  
CC COULD BE INVOLVED IN MYELIN BIOGENESIS AND/OR MYELIN FUNCTION.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO BRAIN AND KIDNEY.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS UP-REGULATED IN  
CC OLIGODENDROCYTES AND BRAIN DURING THE PERIOD OF ACTIVE  
CC MYELINATION.  
CC -!- PTM: LIPOPROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE MAL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U31367; G914968; -.  
DR

DR EMBL; X82557; G854417; -.  
KW TRANSMEMBRANE; LIPOPROTEIN; MYELIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16758 MW; B42E4639 CRC32;  
  
Query Match 35.0%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.97e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 3 PAAASGG 9  
QY 14 PAAASGG 20  
  
RESULT 4  
ID MAL\_CANFA STANDARD; PRT; 153 AA.  
AC Q28296;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN (VIP17 PROTEOLIPID).  
GN MAL.  
OS CANIS FAMILIARIS (DOG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC CARNIVORA; FISSIPEDIA; CANIDAE; CANIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY;  
RX MEDLINE; 96140569.  
RA ZACCHETTI D., PERAENEN J., MURATA M., FIEDLER K., SIMONS K.;  
RT "VIP17/MAL, a proteolipid in apical transport vesicles.";  
RL FEBS LETT. 377:465-469(1995).  
CC -!- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING  
CC -!- CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- PTM: LIPOPROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE MAL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X92505; G1054713; -.  
KW TRANSMEMBRANE; LIPOPROTEIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16624 MW; 9308C3DB CRC32;  
  
Query Match 35.0%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 1.97e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 3 PAAASGG 9  
QY 14 PAAASGG 20  
  
RESULT 5  
ID RL2\_ALTAL STANDARD; PRT; 113 AA.  
AC P42037; P78982;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE 60S ACIDIC RIBOSOMAL PROTEIN P2 (MINOR ALLERGEN ALT A 6) (ALT A VI).  
DE

```
GN ALTA6.
OS ALTERNARIA ALTERNATA.
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; LOCULOASCOMYCETES;
OC MITOSPORIC LOCULOASCOMYCETES; ALTERNARIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-08-0203-BERLIN;
RX MEDLINE; 95206305.
RA ACHATZ G., OBERKOFER H., LECHENAUER E., SIMON B., UNGER A.,
RA KANDLER D., EBNER C., PRILLINGER H., KRAFT D., BREITENBACH M.;
RT "Molecular cloning of major and minor allergens of Alternaria
RT alternata and Cladosporium herbarum.";
RL MOL. IMMUNOL. 32:213-227(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA DE VOUGE M.W., THAKER A.J., ZHANG L., MURADIA G., RODE H.,
RA VIJAY H.M.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- PM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X78222; G467617; -
DR EMBL; U87806; G1850540; -
DR PFAM; PF00428; 60s.Ribosomal; 1.
KW RIBOSOMAL PROTEIN; PHOSPHORYLATION; ALLERGEN.
FT CONFLICT 59 59 S -> P (IN REF. 2).
FT CONFLICT 87 87 Q -> E (IN REF. 2).
SQ SEQUENCE 113 AA; 11134 MW; A2AB209F CRC32;

Query Match 30.0%; Score 6; DB 1; Length 113;
Best Local Similarity 100.0%; Pred.No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 73 AAASGG 78
| | | | |
Qy 15 AAASGG 20

RESULT 6
ID GALA_HUMAN STANDARD; PRT; 123 AA.
AC P22466; Q14413;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE GALANIN PRECURSOR.
GN GALN OR GLNN OR GALL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE OF 1-45 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE; 94140342.
RA EVANS H., BAUMGARTNER M., SHINE J., HERZOG H.;
RT "Genomic organization and localization of the gene encoding human
RT preprogalanin.";
RL GENOMICS 18:473-477(1993).
RN [2]
RP SEQUENCE OF 16-123 FROM N.A.
RX MEDLINE; 92090542.
RA MCKNIGHT G.L., KARLSEN A.E., KOWALYK S., MATHEWES S.L., SHEPPARD P.O.,
```

```
RA O'HARA P.J., TABORSKY G.L.;
RT "Sequence of human galanin and its inhibition of glucose-stimulated
RT insulin secretion from RIN cells.";
RL DIABETES 41:82-87(1992).
RN [3]
RP SEQUENCE OF 33-62.
RX MEDLINE; 91257299.
RA BERSANI M., JOHNSON A.H., HOEJRUP P., DUNNING B.E., ANDREASEN J.J.,
RA HOLST J.J.;
RT "Human galanin: primary structure and identification of two molecular
RT forms.";
RL FEBS LETT. 283:189-194(1991).
RN [4]
RP SEQUENCE OF 33-62, AND SYNTHESIS.
RC TISSUE=PITUITARY;
RX MEDLINE; 92107965.
RA SCHMIDT W.E., KRAYZIN H., ECKART K., DREVS D., MUNDKOWSKI G.,
RA CLEMENS A., KATSOULIS S., SCHAEFER H., GALLWITZ B., CREUTZFELDT W.;
RT "Isolation and primary structure of pituitary human galanin, a 30-
RT residue nonamidated neuropeptide.";
RL PROC. NATL. ACAD. SCI. U.S.A. 88:11435-11439(1991).
CC -1- FUNCTION: CONTRACTS SMOOTH MUSCLE OF THE GASTROINTESTINAL AND
CC GENITOURINARY TRACT, REGULATES GROWTH HORMONE RELEASE, MODULATES
CC INSULIN RELEASE, AND MAY BE INVOLVED IN THE CONTROL OF ADRENAL
CC SECRETION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11144; G488849; -
DR EMBL; M77140; G190412; -
DR PIR; A41630; RHUN.
DR PIR; S15831; S15831.
DR PIR; A49154; A49154.
DR MM; I37035; -
DR PROSITE; PS00861; GALANIN; 1.
DR PFAM; PF01296; Galanin; 1.
KW HORMONE; NEUROPEPTIDE; CLEAVAGE ON PAIR OF BASIC RESIDUES; SIGNAL.
FT SIGNAL 1 19 OR 23 (POTENTIAL).
FT PEPTIDE 33 62 GALANIN.
FT PEPTIDE 65 123 GALANIN MESSAGE-ASSOCIATED PEPTIDE
(GMAP).
SQ SEQUENCE 123 AA; 13302 MW; 38A08263 CRC32;

Query Match 30.0%; Score 6; DB 1; Length 123;
Best Local Similarity 100.0%; Pred.No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 111 LPAAAS 116
| | | | |
Qy 13 LPAAAS 18

RESULT 7
ID RL23_TRYCR STANDARD; PRT; 141 AA.
AC Q94776;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE 60S RIBOSOMAL PROTEIN L23 (L17) (TCSE082).
GN RPL23.
OS TRYPA NOSOMA CRUZI.
OC EUKARYOTA; EULENZOEA; KINETOPLASTIDA; TRYPA NOSOMATIDAE; TRYPA NOSOMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA TANAKA T., TANAKA M.;
```

RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE L14P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D87216; G1536839; -  
DR PROSITE: PS00049; RIBOSOMAL\_L14; 1.  
DR PFAM: PF00238; RIBOSOMAL\_L14; 1.  
DR HSP: P04450; 1WHI.  
KW RIBOSOMAL PROTEIN.  
SQ SEQUENCE 141 AA; 15027 MW; 418BD749 CRC32;  
  
Query Match 30.0%; Score 6; DB 1; Length 141;  
Best Local Similarity 100.0%; Pred. No. 8.80e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 50 RLPA44 55  
QY 12 RLPA44 17  
  
RESULT 8  
ID FMPL\_PSEAE STANDARD; PRT; 157 AA.  
AC P17838;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE FIMBRIAL PROTEIN PRECURSOR (PILIN) (STRAIN PL).  
GN PILA OR FINA.  
OS PSEUDOMONAS AERUGINOSA.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;  
CC PSEUDOMONAS.  
CC [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Pl.  
RX MEDLINE: 88298689.  
RA PASLOSKE B.L., SASTRY P.A., FINLAY B.B., PARANCHYCH W.;  
RT "Two unusual pilin sequences from different isolates of Pseudomonas  
aeruginosa";  
RL J. BACTERIOL. 170:3738-3741(1988).  
CC -!- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4  
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY  
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL  
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.  
CC -!- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M21651; G151474; -  
DR PIR: A31105; A31105.  
DR PROSITE: PS00409; PROKAR\_NTER\_METHYL; 1.  
DR PFAM: PF00114; Pilin; 1.  
DR HSP: P02974; 1AY2.  
KW FIMBRIA; METHYLATION.  
FT PROPEP 1  
FT CHAIN 8 157 FIMBRIAL PROTEIN.  
FT MOD\_RES 8 8 METHYLATION (BY SIMILARITY).  
FT DISULFID 136 149 BY SIMILARITY.  
SQ SEQUENCE 157 AA; 16175 MW; 937B357C CRC32;  
  
Query Match 30.0%; Score 6; DB 1; Length 157;

Best Local Similarity 100.0%; Pred. No. 8.80e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 94 AAASGG 99  
QY 15 AAASGG 20  
  
RESULT 9  
ID ATPD\_HUMAN STANDARD; PRT; 168 AA.  
AC P30049;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE ATP SYNTHASE DELTA CHAIN, MITOCHONDRIAL PRECURSOR (EC 3.6.1.34).  
GN ATP5D.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
CC [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 92182007.  
RA JORDAN E.M., BREEN G.A.M.;  
RT "Molecular cloning of an import precursor of the delta-subunit of the  
human mitochondrial ATP synthase complex";  
RL BIOCHIM. BIOPHYS. ACTA 1130:123-126(1992).  
CC [2]  
RP SEQUENCE FROM N.A.  
RA LAMERDIN J.E., MCCREADY P.M., SKOWRONSKI E., ADAMSON A.W.,  
RA BURKHART-SCHULTZ K., GORDON L., KYLE A., RAMIREZ M., STILWAGEN S.,  
RA PHAN H., VELASCO N., GARNES J., DANGANAN L., POUNDSTONE P.,  
RA CHRISTENSEN M., GEORGESCU A., AVILA J., LIU S., ATTIX C., ANDREISE T.,  
RA TRANKHEIM M., AMICO-KELLER G., COFIELD J., DUARTE S., LUCAS S.,  
RA BRUCE R., THOMAS P., QUAN G., KRONMILLER B., ARELLANO A.,  
RA MONTGOMERY M., OW D., NOLAN M., TRONG S., KOBAYASHI A., OLSEN A.O.,  
RA CARRANO A.V.;  
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC [3]  
RP SEQUENCE OF 23-38.  
RC TISSUE=LIVER;  
RX MEDLINE: 93162045.  
RA HOCHSTRASSER D.F., FRUTIGER S., PAQUET N., BAIRICH A., RAVIER F.,  
RA PASQUALI C., SANCHEZ J.-C., TISSOT J.-D., BUELLOVIST B., VARGAS R.,  
RA APPEL R.D., HUGHES G.J.;  
RT "Human liver protein map: a reference database established by  
microsequencing and gel comparison";  
RL ELECTROPHORESIS 13:992-1001(1992).  
CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
CC GRADIENT ACROSS THE MEMBRANE.  
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
CC SEEMS TO HAVE NINE SUBUNITS: A, B, C, D, E, F, G, F6 AND 8 (OR  
CC A6L).  
CC -!- SIMILARITY: BELONGS TO THE ATPASE EPSILON CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X63422; G12586; -  
DR EMBL: X63423; G12586; -  
DR EMBL: AC004221; G2911258; -  
DR MIM: 603150; -  
DR PIR: S22348; S22348.  
DR SWISS-2DPAGE: P30049; HUMAN.  
DR PFAM: PF00401; ATP-Synt\_DE; 1.  
KW ATP SYNTHESIS; CF(1); HYDROGEN ION TRANSPORT; HYDROLASE;  
KW MITOCHONDRION; TRANSIT PEPTIDE.

```
FT TRANSIT 1 22 MITOCHONDRION.
FT CHAIN 23 168 ATPASE DELTA CHAIN.
SQ SEQUENCE 168 AA; 17490 MW; 0182AE71 CRC32;

Query Match 30.0%; Score 6; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 29 PAAASG 34
   |||||
Qy 14 PAAASG 19

RESULT 10
ID UL45_HSV11 STANDARD; PRT; 172 AA.
AC P10229;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN UL45.
GN UL45
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88274327.
RA MCGEOCH D.J., DALRYMPLE M.A., DAVISON A.J., DOLAN A., FRAME M.C.,
RA MCNAB D., PERRY L.J., SCOTT J.E., TAYLOR P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. GEN. VIROL. 69:1531-1574(1988).
CC -!- SIMILARITY: BELONGS TO THE UL45 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; D10879; G221766; -
DR EMBL; X14112; G59545; -
DR EMBL; X14112; E312360; -
DR PIR; I30088; WMBEF5.
SQ SEQUENCE 172 AA; 18180 MW; 7749AC54 CRC32;

Query Match 30.0%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 RLPAAS 28
   |||||
Qy 12 RLPAAS 17

RESULT 11
ID UL45_HSV1K STANDARD; PRT; 172 AA.
AC P28987;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN UL45 (18 KD PROTEIN).
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN KOS).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 83164308.
RA FRANK R.J., EISENBERG R., COHEN G., WAGNER E.K.;
RT "Detailed analysis of the portion of the herpes simplex virus type 1
RT genome encoding glycoprotein C.";

J. VIROL. 45:634-647(1983).
-!- SIMILARITY: BELONGS TO THE UL45 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; J02216; G330094; -
DR EMBL; K03351; G330097; -
SQ SEQUENCE 172 AA; 18233 MW; 99DD97F3 CRC32;

Query Match 30.0%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 RLPAAS 28
   |||||
Qy 12 RLPAAS 17

RESULT 12
ID UL45_HSV1M STANDARD; PRT; 172 AA.
AC P06482;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN UL45 (18 KD PROTEIN).
GN UL45
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN MP).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84292431.
RA DRAPER K.G., COSTA R.H., LEE G.T.-Y., SPEAR P.G., WAGNER E.K.;
RT "Molecular basis of the glycoprotein-C-negative phenotype of herpes
RT simplex virus type 1 macroplaque strain.";
RL J. VIROL. 51:578-585(1984).
CC -!- SIMILARITY: BELONGS TO THE UL45 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; K03351; G330097; -
DR EMBL; K03351; G330097; -
SQ SEQUENCE 172 AA; 18201 MW; 74E03F63 CRC32;

Query Match 30.0%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 8.80e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 RLPAAS 28
   |||||
Qy 12 RLPAAS 17

RESULT 13
ID CH19_DROGR STANDARD; PRT; 196 AA.
AC P13427;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CHORION PROTEIN S19.
GN CP19 OR S19.
OS DROSOPHILA GRIMSHAWI (FRUIT FLY) (IDIOMYIA GRIMSHAWI).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
```

CC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 88297142.  
RA MARTINEZ-CRUZADO J.C., SWIMMER C., FENERJIAN M.G., KAFATOS F.C.:  
RT "Evolution of the autosomal chorion locus in Drosophila. I. General  
RT organization of the locus and sequence comparisons of genes s15 and  
RT s19 in evolutionary distant species."  
RL GENETICS 119:663-677(1988).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53422; G7422; -  
DR PIR; S06613; S06613.  
DR FLYBASE; FBgn0012320; Dgri\Cp19.  
KW CHORION  
SQ SEQUENCE 196 AA; 19826 MW; 135544F5 CRC32;  
  
Query Match 30.0%; Score 6; DB 1; Length 196;  
Best Local Similarity 100.0%; Pred. No. 8.80e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 56 AAASGG 61  
|||||  
QY 15 AAASGG 20  
  
RESULT 14  
ID NUGM\_PAPR STANDARD; PRT; 209 AA.  
AC P15601;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE PROBABLE NADH-UBIQUINONE OXIDOREDUCTASE SUBUNIT (EC 1.6.5.3) (PROTEIN  
DE P1).  
OS PARAMECIUM PRIMAURELIA.  
OC MITOCHONDRION.  
CC EUKARYOTA; ALVEOLATA; CILIOPHORA; NASSOPHOREA; PENICULIDA; PARAMECIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 8705241.  
RA PRITCHARD A.E., SEILHAMER J.J., CUMMINGS D.J.:  
RT "Paramesium mitochondrial DNA sequences and RNA transcripts for  
RT cytochrome oxidase subunit I, URF1, and three ORFs adjacent to the  
RT replication origin."  
RL GENE 44:243-253(1986).  
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
CC TO BE UBIQUINONE.  
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.  
CC -!- SUBCELLULAR LOCATION: MATRIX AND CYTOPLASMIC SIDE OF THE  
CC MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 30 KD SUBUNIT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M15276; G903855; -  
DR PROSITE; PS00542; COMPLEX1\_30Kd; 1.  
DR PFAM; PF00329; complex1\_30Kd; 1.

KW OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION.  
SQ SEQUENCE 209 AA; 23655 MW; FIA71EDI CRC32;  
  
Query Match 30.0%; Score 6; DB 1; Length 209;  
Best Local Similarity 100.0%; Pred. No. 8.80e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 58 LPAAAS 63  
|||||  
QY 13 LPAAAS 18  
  
RESULT 15  
ID YAAF\_BACSU STANDARD; PRT; 217 AA.  
AC P37529;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 25.4 KD PROTEIN IN SERS-DNAZ INTERGENIC REGION.  
GN YAAF.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
CC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE: 96051385.  
RA OGASAWARA N., NAKAI S., YOSHIKAWA H.;  
RT "Systematic sequencing of the 180 kilobase region of the Bacillus  
RT subtilis chromosome containing the replication origin."  
RL DNA RES. 1:1-14(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D26185; G467404; -  
DR EMBL; Z99104; E1181947; -  
DR SUBTILIS; BG10078; YAAF.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 217 AA; 25444 MW; 630ABA17 CRC32;  
  
Query Match 30.0%; Score 6; DB 1; Length 217;  
Best Local Similarity 100.0%; Pred. No. 8.80e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 28 TLAKRL 33  
|||||  
QY 8 TLAKRL 13

Search completed: Sat Aug 28 14:47:11 1999  
Job time : 47 secs.

**This Page Blank (uspto)**



\*\*\*\*\*  
WISORAH (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:47:30 1999; MasPar time 7.26 Seconds  
Tabular output not generated.  
150.434 Million cell updates/sec

Title: >US-09-049-696-44  
Description: (1-20) from US09049696.pep  
Perfect Score: 20  
Sequence: 1 NSGSDRDLAKRLPAAASGG 20

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: sptrembl9  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 2.719; Variance 0.487; scale 5.585

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	7	35.0	030390	PUTATIVE GONOCOCCAL SE	7.74e-01
2	7	35.0	023004	F6P23.11 PROTEIN.	7.74e-01
3	6	30.0	90 10	PROTEIN KINASE C INHIB	2.53e+01
4	6	30.0	99 2	VACUOLAR TYPE ATP SYNT	2.53e+01
5	6	30.0	107 5	TCP2BETA (HJ.3 LOCUS),	2.53e+01
6	6	30.0	115 2	RTN (FRAGMENT).	2.53e+01
7	6	30.0	120 10	ACIDIC RIBOSOMAL PROTE	2.53e+01
8	6	30.0	132 11	PEX1=HOMEBOX PROTO-ON	2.53e+01
9	6	30.0	155 5	SERINE PROTEINASE (FRA	2.53e+01
10	6	30.0	162 6	CAVEOLIN-2.	2.53e+01
11	6	30.0	202 4	HYPOTHETICAL 21.7 KD P	2.53e+01
12	6	30.0	203 14	COAT PROTEIN.	2.53e+01
13	6	30.0	204 3	HYPOTHETICAL 22.6 KD P	2.53e+01
14	6	30.0	206 5	T22H6.1 PROTEIN.	2.53e+01
15	6	30.0	209 5	BLASTODERM-SPECIFIC PR	2.53e+01
16	6	30.0	214 2	HYPOTHETICAL 23.7 KD P	2.53e+01
17	6	30.0	218 2	TRANSFERRIN BINDING PR	2.53e+01
18	6	30.0	227 6	NEURONAL AXONAL MEMBR	2.53e+01
19	6	30.0	245 10	ACC SYNTHASE (EC 4.1.1	2.53e+01
20	6	30.0	245 4	WUGSC:H_DJ077023.2 PR	2.53e+01

WINGLESS (FRAGMENT). 2.53e+01  
MYOSIN HEAVY CHAIN (FR 2.53e+01  
CHLOROPHYLL A/B BINDIN 2.53e+01  
LIGHT HARVESTING CHLO 2.53e+01  
HEPATIC GLYCOGEN-BINDI 2.53e+01  
HYPOTHETICAL 32.3 KD P 2.53e+01  
NITROGENASE REDUCTASE. 2.53e+01  
CALDENDRIN. 2.53e+01  
L1439.3. 2.53e+01  
RNA POLYMERASE SIGMA F 2.53e+01  
HYPOTHETICAL 36.5 KD P 2.53e+01  
COSMID F28B3. 2.53e+01  
GLYCOPROTEIN E (FRAGME 2.53e+01  
PUTATIVE TWO-COMPONENT 2.53e+01  
HYPOTHETICAL 44.9 KD P 2.53e+01  
F8K4.10 PROTEIN. 2.53e+01  
PBX3A. 2.53e+01  
HYPOTHETICAL 48.2 KD P 2.53e+01  
RIBULOSE BISPHOSPHATE 2.53e+01  
HYPOTHETICAL 50.6 KD P 2.53e+01  
UL38. 2.53e+01  
O-ANTIGEN TRANSPORTER 2.53e+01  
FLAGELLIN (FRAGMENT). 2.53e+01  
TUBBY PROTEIN. 2.53e+01  
PUTATIVE TRANSPORT PRO 2.53e+01  
GLYCOPROTEIN E. 2.53e+01  
GLYCOPROTEIN GE PRECUR 2.53e+01  
CDPK-RELATED PROTEIN K 2.53e+01  
NUMBLIKE. 2.53e+01  
CDPK-RELATED PROTEIN K 2.53e+01  
T7N9.22. 2.53e+01  
CALCIUM/CALMODULIN-DEP 2.53e+01  
POLYADENYLATE BINDING 2.53e+01  
F48A11.5 PROTEIN. 2.53e+01  
TRANSFERIN BINDING PRO 2.53e+01  
TRANSFERIN-BINDING PR 2.53e+01  
TRANSFERIN-BINDING PR 2.53e+01  
(S3032) TBP2 GENE (FRA 2.53e+01  
TRANSFERIN BINDING PRO 2.53e+01  
(M978) TBP2 GENE (FRAG 2.53e+01  
TONB-DEPENDENT RECEPTO 2.53e+01  
HYPOTHETICAL 74.5 KD P 2.53e+01  
INSOLUBLE PROTEIN. 2.53e+01  
RNA HELICASE ISOLOGPH. 2.53e+01  
VACUOLAR H+-PYROPHOSPH 2.53e+01  
IRON-REGULATED OUTER M 2.53e+01  
CDS LOCALIZED AFTER CO 2.53e+01  
PROTEIN ASSOCIATED WIT 2.53e+01  
LEVANASE (EC 3.2.1.26) 2.53e+01  
ATROPHIN-1 RELATED PRO 2.53e+01  
ATROPHIN-1 LIKE PROTEI 2.53e+01  
HYPOTHETICAL 115.0 KD 2.53e+01  
CTD-BINDING SR-LIKE PR 2.53e+01  
BRAHMA ASSOCIATED PROT 2.53e+01  
HYPOTHETICAL 138.8 KD 2.53e+01  
KIAA0335. 2.53e+01  
KIAA0458 PROTEIN. 2.53e+01  
L1439.2. 2.53e+01  
MYOSIN. 2.53e+01  
X GENE. 2.53e+01  
POLYKETIDE SYNTHASE. 2.53e+01  
RIFAMYCIN POLYKETIDE S 2.53e+01  
PEXP59 PROTEIN (FRAGME 5.80e+02  
HOMEBOX PROTEIN LPHOX 5.80e+02  
R2R3-MYB TRANSCRIPTION 5.80e+02  
GLYCOPROTEIN GD (FRAGM 5.80e+02  
PUTATIVE GLUTAMATE DEC 5.80e+02  
ENVELOPE PROTEIN (FRAG 5.80e+02  
COSMID C32B8. 5.80e+02  
REARRANGED MLL PROTEIN 5.80e+02  
VPF77. 5.80e+02  
HYPOTHETICAL PROTEIN H 5.80e+02  
CALMODULIN-DEPENDENT P 5.80e+02

94	5	25.0	80 10	O22488	METALLOTHIONEIN-LIKE P	5.80e+02	167	5	25.0	139 2	O51862	GYRASE B SUBUNIT (FRAG	5.80e+02
95	5	25.0	80 14	Q83874	HELICASE (FRAGMENT)	5.80e+02	168	5	25.0	139 2	O59764	NICOTINAMIDE NUCLEOTID	5.80e+02
96	5	25.0	82 7	Q30544	MHC CLASS II BETA CHAI	5.80e+02	169	5	25.0	140 14	O72499	POLYPROTEIN (FRAGMENT)	5.80e+02
97	5	25.0	83 7	Q30543	MHC CLASS II BETA CHAI	5.80e+02	170	5	25.0	140 14	O72500	POLYPROTEIN (FRAGMENT)	5.80e+02
98	5	25.0	83 7	Q30547	MHC CLASS II BETA CHAI	5.80e+02	171	5	25.0	140 14	O72498	POLYPROTEIN (FRAGMENT)	5.80e+02
99	5	25.0	85 2	Q32905	ANTIGEN 43L (FRAGMENT)	5.80e+02	172	5	25.0	140 14	O88927	HELPER COMPONENT-RELAT	5.80e+02
100	5	25.0	87 2	P94353	HYPOTHETICAL PROTEIN.	5.80e+02	173	5	25.0	140 14	O72495	POLYPROTEIN (FRAGMENT)	5.80e+02
101	5	25.0	92 6	O29281	UNKNOWN PROTEIN (FRAGM	5.80e+02	174	5	25.0	140 5	O18318	HISTONE H1-II-2 (FRAG	5.80e+02
102	5	25.0	94 11	O08937	PUTATIVE ADRENOMEDULLI	5.80e+02	175	5	25.0	141 2	O67689	HYPOTHETICAL 15.9 KD P	5.80e+02
103	5	25.0	95 11	O64086	OR-OXYTOCIN RECEPTOR	5.80e+02	176	5	25.0	141 6	O79297	RETINOIC ACID RECEPTOR	5.80e+02
104	5	25.0	96 3	O42658	HYPOTHETICAL 11.1 KD P	5.80e+02	177	5	25.0	141 14	O72501	POLYPROTEIN (FRAGMENT)	5.80e+02
105	5	25.0	100 9	Q38067	ORF 100.	5.80e+02	178	5	25.0	141 14	O72503	POLYPROTEIN (FRAGMENT)	5.80e+02
106	5	25.0	101 10	O43059	1-AMINOCYCLOPROPANE-1-	5.80e+02	179	5	25.0	143 14	O71110	ADENOVIRUS 3 COMPLETE	5.80e+02
107	5	25.0	101 10	O04898	1-AMINOCYCLOPROPANE-1-	5.80e+02	180	5	25.0	143 2	O47548	DAPX PROTEIN.	5.80e+02
108	5	25.0	105 2	O07632	YL4H PROTEIN.	5.80e+02	181	5	25.0	144 2	P78000	KDPB.	5.80e+02
109	5	25.0	106 2	O49910	L308_C3_265.	5.80e+02	182	5	25.0	144 5	O00878	MEROZOITE SURFACE PROT	5.80e+02
110	5	25.0	112 14	O56586	GLYCOPROTEIN 120 (FRAG	5.80e+02	183	5	25.0	147 13	Q12940	TOM-1A PROTEIN.	5.80e+02
111	5	25.0	113 2	O87067	Z61R PROTEIN.	5.80e+02	184	5	25.0	149 3	Q12258	ORF YL052C.	5.80e+02
112	5	25.0	119 14	O89165	POLYPROTEIN (FRAGMENT)	5.80e+02	185	5	25.0	150 2	O56269	SERINE-RICH PROTEIN (F	5.80e+02
113	5	25.0	119 2	O48980	50S RIBOSOMAL PROTEIN	5.80e+02	186	5	25.0	152 2	O45956	ORF 152B.	5.80e+02
114	5	25.0	120 6	O29275	URIDYLATE KINASE (FRAG	5.80e+02	187	5	25.0	152 2	O52895	HYPOTHETICAL 17.5 KD P	5.80e+02
115	5	25.0	124 2	O53292	HYPOTHETICAL 13.4 KD P	5.80e+02	188	5	25.0	153 5	O00879	MEROZOITE SURFACE PROT	5.80e+02
116	5	25.0	125 2	O09034	HYPOTHETICAL 14.6 KD P	5.80e+02	189	5	25.0	154 2	P71755	HYPOTHETICAL 15.7 KD P	5.80e+02
117	5	25.0	125 6	O77766	M3 MUSCARINIC ACETYLCH	5.80e+02	190	5	25.0	157 2	P97029	HYPOTHETICAL 17.8 KD P	5.80e+02
118	5	25.0	125 2	O47664	POT. ORF (FRAGMENT).	5.80e+02	191	5	25.0	157 1	O28464	INOSINE MONOPHOSPHATE	5.80e+02
119	5	25.0	131 4	O43612	PREPRO-OREXIN.	5.80e+02	192	5	25.0	158 2	O54650	FIMBRIAL SUBUNIT PROTE	5.80e+02
120	5	25.0	131 14	O66235	POLYPROTEIN.	5.80e+02	193	5	25.0	159 3	P79041	COSMID SM06E08; HTGS P	5.80e+02
121	5	25.0	133 14	O41696	ENVELOPE PROTEIN 2 (FR	5.80e+02	194	5	25.0	159 2	O69055	PUTATIVE C-P LYASE SUB	5.80e+02
122	5	25.0	133 14	O41695	ENVELOPE PROTEIN 2 (FR	5.80e+02	195	5	25.0	160 2	Q03542	TRHB PROTEIN.	5.80e+02
123	5	25.0	133 14	O41694	ENVELOPE PROTEIN 2 (FR	5.80e+02	196	5	25.0	162 2	P95815	CYTCHROME P-450-9-DEO	5.80e+02
124	5	25.0	133 14	O41733	ENVELOPE PROTEIN 2 (FR	5.80e+02	197	5	25.0	162 2	P71183	TRHB.	5.80e+02
125	5	25.0	133 14	O41723	ENVELOPE PROTEIN 2 (FR	5.80e+02	198	5	25.0	162 1	O50385	METHYL-COENZYME M REDU	5.80e+02
126	5	25.0	133 14	O41709	ENVELOPE PROTEIN 2 (FR	5.80e+02	199	5	25.0	163 1	O50384	METHYL-COENZYME M REDU	5.80e+02
127	5	25.0	133 14	O41708	ENVELOPE PROTEIN 2 (FR	5.80e+02	200	5	25.0	163 1	O50216	METHYL-COENZYME M REDU	5.80e+02
128	5	25.0	133 14	O41690	ENVELOPE PROTEIN 2 (FR	5.80e+02	201	5	25.0	163 5	Q18036	SIMILAR TO NADH-UBIQUI	5.80e+02
129	5	25.0	133 14	O41691	ENVELOPE PROTEIN 2 (FR	5.80e+02	202	5	25.0	163 2	O53000	CSHD PROTEIN.	5.80e+02
130	5	25.0	133 14	O41687	ENVELOPE PROTEIN 2 (FR	5.80e+02	203	5	25.0	163 1	O48889	METHYL-COENZYME M REDU	5.80e+02
131	5	25.0	133 14	O41685	ENVELOPE PROTEIN 2 (FR	5.80e+02	204	5	25.0	166 5	O93533	F20D1.1 PROTEIN.	5.80e+02
132	5	25.0	133 14	O41728	ENVELOPE PROTEIN 2 (FR	5.80e+02	205	5	25.0	167 10	O40232	MANNOSE-BINDING PROTEI	5.80e+02
133	5	25.0	133 14	O41721	ENVELOPE PROTEIN 2 (FR	5.80e+02	206	5	25.0	168 10	O39729	LECTIN PRECURSOR.	5.80e+02
134	5	25.0	133 14	O41739	ENVELOPE PROTEIN 2 (FR	5.80e+02	207	5	25.0	168 10	O65051	PUTATIVE RNA-BINDING P	5.80e+02
135	5	25.0	133 14	O41706	ENVELOPE PROTEIN 2 (FR	5.80e+02	208	5	25.0	172 1	O58522	172AA LONG HYPOTHETICA	5.80e+02
136	5	25.0	133 14	O41688	ENVELOPE PROTEIN 2 (FR	5.80e+02	209	5	25.0	172 2	O51670	HYPOTHETICAL 18.8 KD P	5.80e+02
137	5	25.0	133 14	O41689	ENVELOPE PROTEIN 2 (FR	5.80e+02	210	5	25.0	174 2	P96943	OUTER MEMBRANE PROTEIN	5.80e+02
138	5	25.0	133 14	O41700	ENVELOPE PROTEIN 2 (FR	5.80e+02	211	5	25.0	174 2	P95372	OUTER MEMBRANE PROTEIN	5.80e+02
139	5	25.0	133 14	O41684	ENVELOPE PROTEIN 2 (FR	5.80e+02	212	5	25.0	176 14	O89784	POTENTIAL SIGNAL PEPTI	5.80e+02
140	5	25.0	133 14	O41702	ENVELOPE PROTEIN 2 (FR	5.80e+02	213	5	25.0	176 2	O07747	HYPOTHETICAL 18.0 KD P	5.80e+02
141	5	25.0	133 14	O41735	ENVELOPE PROTEIN 2 (FR	5.80e+02	214	5	25.0	178 1	O29147	HYPOTHETICAL 19.3 KD P	5.80e+02
142	5	25.0	133 14	O41719	ENVELOPE PROTEIN 2 (FR	5.80e+02	215	5	25.0	179 5	O94732	SALIVARY PLATELET AGGR	5.80e+02
143	5	25.0	133 14	O41720	ENVELOPE PROTEIN 2 (FR	5.80e+02	216	5	25.0	179 14	O66667	CAPSID PROTEIN.	5.80e+02
144	5	25.0	133 14	O41703	ENVELOPE PROTEIN 2 (FR	5.80e+02	217	5	25.0	179 14	O64889	MAJOR CORE PROTEIN.	5.80e+02
145	5	25.0	133 14	O41681	ENVELOPE PROTEIN 2 (FR	5.80e+02	218	5	25.0	179 14	O65194	STRUCTURAL PROTEIN P54	5.80e+02
146	5	25.0	133 14	O41680	ENVELOPE PROTEIN 2 (FR	5.80e+02	219	5	25.0	183 14	O65272	J13L GENE (ISOLATE LIS	5.80e+02
147	5	25.0	133 10	O50508	60S RIBOSOMAL PROTEIN	5.80e+02	220	5	25.0	183 14	O65272	J13L GENE (ISOLATE LIS	5.80e+02
148	5	25.0	133 14	O41686	ENVELOPE PROTEIN 2 (FR	5.80e+02	221	5	25.0	184 14	O98144	GENE 20 PROTEIN (FRAGM	5.80e+02
149	5	25.0	133 14	O41697	ENVELOPE PROTEIN 2 (FR	5.80e+02	222	5	25.0	185 14	O65271	J13L GENE (ISOLATE LIS	5.80e+02
150	5	25.0	133 14	O41682	ENVELOPE PROTEIN 2 (FR	5.80e+02	223	5	25.0	189 14	O65270	J13L GENE (ISOLATE LIS	5.80e+02
151	5	25.0	133 14	O41729	ENVELOPE PROTEIN 2 (FR	5.80e+02	224	5	25.0	189 14	O64891	E1A ORF1.	5.80e+02
152	5	25.0	133 14	O41730	ENVELOPE PROTEIN 2 (FR	5.80e+02	225	5	25.0	191 2	O85349	STSL.	5.80e+02
153	5	25.0	133 14	O41725	ENVELOPE PROTEIN 2 (FR	5.80e+02	226	5	25.0	191 5	O45682	K10H10.5 PROTEIN.	5.80e+02
154	5	25.0	133 14	O41727	ENVELOPE PROTEIN 2 (FR	5.80e+02	227	5	25.0	193 2	O04793	ERYBVII (ORF19).	5.80e+02
155	5	25.0	133 14	O41692	ENVELOPE PROTEIN 2 (FR	5.80e+02	228	5	25.0	194 14	O84196	22K PROTEIN.	5.80e+02
156	5	25.0	133 14	O41693	ENVELOPE PROTEIN 2 (FR	5.80e+02	229	5	25.0	194 2	P74480	HYPOTHETICAL 20.8 KD P	5.80e+02
157	5	25.0	133 14	O41699	ENVELOPE PROTEIN 2 (FR	5.80e+02	230	5	25.0	194 2	O86619	HYPOTHETICAL 22.1 KD P	5.80e+02
158	5	25.0	133 14	O41683	ENVELOPE PROTEIN 2 (FR	5.80e+02	231	5	25.0	195 2	O08250	BIOTIN TRANSPORT REGUL	5.80e+02
159	5	25.0	135 10	Q40228	ORF (FRAGMENT).	5.80e+02	232	5	25.0	195 14	O42050	MATRIX PROTEIN 2.	5.80e+02
160	5	25.0	136 3	O02608	HISTONE H2B.	5.80e+02	233	5	25.0	195 11	Q01585	MYELIN BASIC PROTEIN (	5.80e+02
161	5	25.0	136 2	O06815	HYPOTHETICAL 15.0 KD P	5.80e+02	234	5	25.0	196 14	O65275	J13L GENE (ISOLATE UGA	5.80e+02
162	5	25.0	136 10	O22674	REVERSE TRANSCRIPTASE	5.80e+02	235	5	25.0	197 5	O93103	PREPROALLATOSTATIN.	5.80e+02
163	5	25.0	137 14	O98492	GENOME. PARTIAL SEQUEN	5.80e+02	236	5	25.0	199 2	O86325	HYPOTHETICAL 22.6 KD P	5.80e+02
164	5	25.0	138 10	P93631	HMG22 PROTEIN.	5.80e+02	237	5	25.0	200 10	O40209	RAB5A.	5.80e+02
165	5	25.0	138 1	O26976	CONSERVED PROTEIN.	5.80e+02	238	5	25.0	200 2	O05326	RAB5B.	5.80e+02
166	5	25.0	139 10	P93630	HMGCI PROTEIN.	5.80e+02	239	5	25.0	201 10	O65842	SUBUNIT B PROTEIN OF D	5.80e+02
												SMALL GTP-BINDING PROT	5.80e+02

240	5	25.0	204	14	065843	5'	END OF ORF 1 UNDETE	5	80e+02	313	5	25.0	281	2	053383	HYPOTHETICAL 28.7 KD P	5	80e+02
241	5	25.0	206	14	011467	CAPSID PROTEIN (FRAGME		5	80e+02	314	5	25.0	282	2	044259	CATECHOL 1,2-DIOXYGENA	5	80e+02
242	5	25.0	207	2	007559	HYPOTHETICAL 23.3 KD P		5	80e+02	315	5	25.0	282	2	031247	CATECHOL 1,2-DIOXYGENA	5	80e+02
243	5	25.0	210	11	070377	SNAP-23.		5	80e+02	316	5	25.0	282	10	034357	HEAT SHOCK TRANSCRIPT I	5	80e+02
244	5	25.0	211	11	009044	SYNDET (SNAP-23).		5	80e+02	317	5	25.0	283	2	067855	PROTEINASE IV.	5	80e+02
245	5	25.0	211	1	029897	CONSERVED HYPOTHETICAL		5	80e+02	318	5	25.0	283	2	P73114	ESTERASE.	5	80e+02
246	5	25.0	215	2	005521	DNA FOR PHOB-RNNE-GROE		5	80e+02	319	5	25.0	283	10	081988	CP31BHV.	5	80e+02
247	5	25.0	216	10	043114	PECTINESTERASE (EC 3.1		5	80e+02	320	5	25.0	284	10	085794	HISTONE H1.	5	80e+02
248	5	25.0	216	2	067529	SOLUBLE HYDROGENASE SM		5	80e+02	321	5	25.0	288	14	036077	COAT PROTEIN.	5	80e+02
249	5	25.0	216	3	074716	ENDO-1,4-BETA-XYLANASE		5	80e+02	322	5	25.0	289	3	P78845	FISSION YEAST (FRAGMEN	5	80e+02
250	5	25.0	217	2	007302	MERCURIC RESISTANCE PR		5	80e+02	323	5	25.0	290	2	011198	HYPOTHETICAL 30.3 KD P	5	80e+02
251	5	25.0	219	14	066870	CAPSID PROTEIN (FRAGME		5	80e+02	324	5	25.0	291	5	006301	SM50 ANTIGEN (FRAGMENT	5	80e+02
252	5	25.0	219	14	068869	CAPSID PROTEIN (FRAGME		5	80e+02	325	5	25.0	292	2	033260	HYPOTHETICAL 31.8 KD P	5	80e+02
253	5	25.0	219	14	068867	CAPSID PROTEIN (FRAGME		5	80e+02	326	5	25.0	293	1	030226	N5-METHYLTETRAHYDROMET	5	80e+02
254	5	25.0	220	5	018313	HOX 5.		5	80e+02	327	5	25.0	293	2	087716	CHL.	5	80e+02
255	5	25.0	220	14	081992	CAPSID PROTEIN L1 (FRA		5	80e+02	328	5	25.0	294	2	005699	GLYCOSYL TRANSFERASE.	5	80e+02
256	5	25.0	221	5	009407	HYPOTHETICAL 26.3 KD P		5	80e+02	329	5	25.0	294	5	076361	F55A8.1 PROTEIN (FRAGM	5	80e+02
257	5	25.0	223	4	P13985	HTLV-1 RELATED ENDOGEN		5	80e+02	330	5	25.0	295	3	012219	YOR3248W FROM CHROMOSO	5	80e+02
258	5	25.0	224	9	003901	ORFS P21, P11.5, P26 &		5	80e+02	331	5	25.0	297	2	093931	HYPOTHETICAL 32 KD PRO	5	80e+02
259	5	25.0	225	14	007785	PN57-2.		5	80e+02	332	5	25.0	297	10	039931	ACC SYNTHASE (FRAGMENT	5	80e+02
260	5	25.0	225	2	044479	ORF225.		5	80e+02	333	5	25.0	298	5	062415	Y22F5A.4 PROTEIN.	5	80e+02
261	5	25.0	226	5	018317	HISTONE H1-II-2.		5	80e+02	334	5	25.0	299	2	056171	N-ACETYLPHOSPHINOTHRIC	5	80e+02
262	5	25.0	227	2	055747	HYPOTHETICAL 24.5 KD P		5	80e+02	335	5	25.0	299	2	004816	ORF9 PROTEIN PRECURSOR	5	80e+02
263	5	25.0	227	3	042621	IMIDAZOLEGLYCEROL-PHOS		5	80e+02	336	5	25.0	299	5	021682	COSMID R03H10.	5	80e+02
264	5	25.0	227	5	023792	HISTONE H1-II-2.		5	80e+02	337	5	25.0	301	4	014799	INACTIVE PALMITOYL-PRO	5	80e+02
265	5	25.0	227	5	018316	HISTONE H1-II-2.		5	80e+02	338	5	25.0	303	1	027947	CONSERVED HYPOTHETICAL	5	80e+02
266	5	25.0	228	14	069344	THYMIDINE KINASE (FRAG		5	80e+02	339	5	25.0	303	2	006545	HYPOTHETICAL 32.5 KD P	5	80e+02
267	5	25.0	229	2	053930	HISD L-HISTIDINOL:NAD+		5	80e+02	340	5	25.0	304	2	045500	HYPOTHETICAL 34.8 KD P	5	80e+02
268	5	25.0	231	10	041499	SPLICEOSOMAL PROTEIN.		5	80e+02	341	5	25.0	305	2	053129	CHEMOTAXIS OPERON DNA.	5	80e+02
269	5	25.0	231	5	023785	HISTONE H1-III.		5	80e+02	342	5	25.0	305	5	045424	F31D4.1 PROTEIN.	5	80e+02
270	5	25.0	232	5	016630	KO2F6.8 PROTEIN.		5	80e+02	343	5	25.0	307	2	049138	(CLONE PDN9, HINDIIIAB	5	80e+02
271	5	25.0	233	2	063783	DAPB.		5	80e+02	344	5	25.0	308	4	099945	HYPOTHETICAL 34.9 KD P	5	80e+02
272	5	25.0	233	5	018319	HISTONE H1D.		5	80e+02	345	5	25.0	308	2	050428	PGRS-FAMILY.	5	80e+02
273	5	25.0	234	2	066914	3-DEOXY-MANNO-OCTULOSO		5	80e+02	346	5	25.0	308	14	039491	SUBSTRATE OF SERINE PR	5	80e+02
274	5	25.0	236	2	053157	UNKNOWN ORF.		5	80e+02	347	5	25.0	308	2	034090	PORPHOBILINOGEN DEAMIN	5	80e+02
275	5	25.0	238	10	023300	CARNITINE RACEMASE HOM		5	80e+02	348	5	25.0	311	10	023381	SIMILARITY TO UTP-GLUC	5	80e+02
276	5	25.0	238	2	053200	HYPOTHETICAL 25.1 KD P		5	80e+02	349	5	25.0	311	3	074848	HYPOTHETICAL 35.7 KD P	5	80e+02
277	5	25.0	238	10	049203	NUCLEOSIDE DIPHOSPHATE		5	80e+02	350	5	25.0	312	1	073945	HYPOTHETICAL 37.7 KD P	5	80e+02
278	5	25.0	242	2	068101	UROPORHYRIN-III C-MET		5	80e+02	351	5	25.0	312	2	033744	HYPOTHETICAL 38.8 KD P	5	80e+02
279	5	25.0	244	5	023972	SEVELIN CLONE D69 CELL		5	80e+02	352	5	25.0	313	2	050268	MOAB.	5	80e+02
280	5	25.0	246	2	P94338	SIMILAR TO YFTH.		5	80e+02	353	5	25.0	313	10	081912	T7133.18 PROTEIN.	5	80e+02
281	5	25.0	251	9	064332	GP18.		5	80e+02	354	5	25.0	314	2	031795	TRNA ISOPENTENYLVPROPH	5	80e+02
282	5	25.0	251	5	017435	MINOR AMPULLATE SILK P		5	80e+02	355	5	25.0	314	10	043561	PROTCHLOROPHYLLIDE RE	5	80e+02
283	5	25.0	253	11	035779	ENHANCER-OF-SPLIT AND		5	80e+02	356	5	25.0	314	5	017243	F53G2.5 PROTEIN.	5	80e+02
284	5	25.0	253	3	042737	NECROSIS-INDUCING PEPT		5	80e+02	357	5	25.0	314	5	045304	C47A10.2 PROTEIN.	5	80e+02
285	5	25.0	254	1	029772	ATP-BINDING PROTEIN PH		5	80e+02	358	5	25.0	316	13	093489	LJOTXB.	5	80e+02
286	5	25.0	256	5	018599	DVTRY-1 TRYPSIN PRECUR		5	80e+02	359	5	25.0	316	2	084535	METHIONYL TRNA FORMYLT	5	80e+02
287	5	25.0	257	10	004140	MYB FACTOR, 1202 BP.		5	80e+02	360	5	25.0	316	10	P93052	MALATE DEHYDROGENASE.	5	80e+02
288	5	25.0	259	2	051348	AMINOACYL-ADENILTRANSF		5	80e+02	361	5	25.0	317	10	022176	PUTATIVE ELICITOR RESP	5	80e+02
289	5	25.0	260	5	076669	H10E21.1 PROTEIN.		5	80e+02	362	5	25.0	317	1	026383	PROTEIN KINASE.	5	80e+02
290	5	25.0	260	2	045038	(CLONE 8).		5	80e+02	363	5	25.0	317	11	061890	RECEPTOR-LIKE TYROSINE	5	80e+02
291	5	25.0	261	2	055715	HYPOTHETICAL 29.0 KD P		5	80e+02	364	5	25.0	318	10	040781	HOMEOBOX-LEUCINE ZIPPE	5	80e+02
292	5	25.0	261	10	038828	AUXIN-INDUCIBLE IAA10.		5	80e+02	365	5	25.0	319	14	065410	POLYPROTEIN (FRAGMENTS	5	80e+02
293	5	25.0	263	2	005346	INTRINSIC MEMBRANE PRO		5	80e+02	366	5	25.0	319	14	065411	POLYPROTEIN (FRAGMENTS	5	80e+02
294	5	25.0	263	2	034859	YOJE PROTEIN.		5	80e+02	367	5	25.0	320	5	P91261	HYPOTHETICAL 35.4 KD P	5	80e+02
295	5	25.0	263	10	048991	NBS-LRR TYPE RESISTANC		5	80e+02	368	5	25.0	322	14	010388	RNA 3.	5	80e+02
296	5	25.0	265	5	016631	KO2F6.5 PROTEIN.		5	80e+02	369	5	25.0	323	2	007843	HYPOTHETICAL 35.0 KD P	5	80e+02
297	5	25.0	265	2	085953	TRANGLYCOSYLASE PRECU		5	80e+02	370	5	25.0	324	2	030449	HYPOTHETICAL 34.3 KD P	5	80e+02
298	5	25.0	270	2	031917	YOOV PROTEIN.		5	80e+02	371	5	25.0	326	5	022379	SIMILAR TO OTHER C. EL	5	80e+02
299	5	25.0	270	2	006130	TRPA.		5	80e+02	372	5	25.0	328	2	056126	OUTER MEMBRANE PROTEIN	5	80e+02
300	5	25.0	270	2	067488	HYPOTHETICAL 31.0 KD P		5	80e+02	373	5	25.0	328	2	056126	CHROMOSOME IV READING	5	80e+02
301	5	25.0	271	2	068906	TRYPTOPHAN SYNTHASE AL		5	80e+02	374	5	25.0	329	14	P90341	HERPES SIMPLEX VIRUS T	5	80e+02
302	5	25.0	272	1	039526	272AA LONG HYPOTHETICA		5	80e+02	375	5	25.0	329	2	052472	L-FUCULOSE DEHYDROGENASE	5	80e+02
303	5	25.0	275	2	033969	LIPASE LIPA PRECURSOR.		5	80e+02	376	5	25.0	329	10	043562	PROTCHLOROPHYLLIDE RE	5	80e+02
304	5	25.0	276	2	066754	ACETYL-COA CARBOXYLTRA		5	80e+02	377	5	25.0	329	10	064640	F17K2.13 PROTEIN.	5	80e+02
305	5	25.0	276	3	036033	HYPOTHETICAL PROTEIN C		5	80e+02	378	5	25.0	330	2	005274	SUBUNIT B PROTEIN OF D	5	80e+02
306	5	25.0	276	10	004108	MYB FACTOR.		5	80e+02	379	5	25.0	330	2	050524	HYPOTHETICAL 34.4 KD P	5	80e+02
307	5	25.0	277	5	015771	MUCIN-LIKE PROTEIN (FR		5	80e+02	380	5	25.0	330	2	005276	SUBUNIT B PROTEIN OF D	5	80e+02
308	5	25.0	278	2	P77959	RNA POLYMERASE SIGMA F		5	80e+02	381	5	25.0	330	2	005275	SUBUNIT B PROTEIN OF D	5	80e+02
309	5	25.0	278	2	006384	RNA POLYMERASE SIGMA F		5	80e+02	382	5	25.0	333	2	030331	HYPOTHETICAL 37.2 KD P	5	80e+02
310	5	25.0	278	2	059823	RNA POLYMERASE SIGMA F		5	80e+02	383	5	25.0	334	1	058417	334AA LONG HYPOTHETICA	5	80e+02
311	5	25.0	281	5	025344	L2759.5 PROTEIN.		5	80e+02	384	5	25.0	335	4	043403	RAD51-INTERACTING PROT	5	80e+02
312	5	25.0	281	2	087598	20S PROTEASOME BETA-SU		5	80e+02	385	5	25.0	335	10	081384	PUTATIVE BASIC HELIX-L	5	80e+02

386	5	25.0	335	2	085011	HYPOTHETICAL 34.4 KD P	5.80e+02	459	5	25.0	385	2	052323	HYPOTHETICAL 41.8 KD P	5.80e+02
387	5	25.0	336	10	080648	PUTATIVE DOLICHYL-PHOS	5.80e+02	460	5	25.0	386	2	087199	HOMOAONITATE HYDRATAS	5.80e+02
388	5	25.0	339	3	074959	HYPOTHETICAL 38.5 KD P	5.80e+02	461	5	25.0	386	4	013746	BCR-ABL MRNA OF ACUTE	5.80e+02
389	5	25.0	340	2	044324	MOCOD.	5.80e+02	462	5	25.0	387	10	038833	PUTATIVE CHLOROPHYLL S	5.80e+02
390	5	25.0	340	2	076142	FROM BASES 1598243 TO	5.80e+02	463	5	25.0	388	1	058092	388AA LONG HYPOTHETICA	5.80e+02
391	5	25.0	342	5	018888	SIMILAR TO TPR DOMAIN	5.80e+02	464	5	25.0	388	2	077977	CYTCHROME P450 (EC 1.	5.80e+02
392	5	25.0	343	2	031569	YFHA PROTEIN.	5.80e+02	465	5	25.0	390	2	050279	PEPB.	5.80e+02
393	5	25.0	345	1	026403	CONSERVED PROTEIN.	5.80e+02	466	5	25.0	390	1	073950	IRON CONTAINING ALCOHO	5.80e+02
394	5	25.0	347	5	024502	SYNAPSE-ASSOCIATED PRO	5.80e+02	467	5	25.0	391	2	073193	SPOULATION PROTEIN SP	5.80e+02
395	5	25.0	349	3	087249	HAPB PROTEIN.	5.80e+02	468	5	25.0	391	4	075910	ECTODYSPLASIN-A,EDA IS	5.80e+02
396	5	25.0	349	2	007146	HYPOTHETICAL 37.9 KD P	5.80e+02	469	5	25.0	392	2	085745	DNA GRASE BETA-SUBUNI	5.80e+02
397	5	25.0	349	10	081756	HYPOTHETICAL 38.4 KD P	5.80e+02	470	5	25.0	392	2	030529	POPB.	5.80e+02
398	5	25.0	351	5	001442	CODED FOR BY C. ELEGAN	5.80e+02	471	5	25.0	393	2	050657	LYSINE/ORNITHINE DECAR	5.80e+02
399	5	25.0	351	5	024503	SYNAPSE ASSOCIATED PRO	5.80e+02	472	5	25.0	394	10	043809	1-AMINOCYCLOPROPANE 1-	5.80e+02
400	5	25.0	351	2	006205	HYPOTHETICAL 38.1 KD P	5.80e+02	473	5	25.0	394	2	052265	CLASS D TETRACYCLINE/H	5.80e+02
401	5	25.0	352	14	039280	COUNTERPART OF HSV-1 G	5.80e+02	474	5	25.0	398	2	077842	RNA POLYMERASE SIGMA F	5.80e+02
402	5	25.0	352	14	066596	MUTANT THYMIDINE KINAS	5.80e+02	475	5	25.0	400	5	026652	NF-Y-A SUBUNIT.	5.80e+02
403	5	25.0	352	2	074819	GLYCOSYL TRANSFERASE.	5.80e+02	476	5	25.0	400	2	045564	TRANSACYLASE.	5.80e+02
404	5	25.0	352	2	024763	MXAR PROTEIN.	5.80e+02	477	5	25.0	403	1	058709	HYPOTHETICAL PROTEIN M	5.80e+02
405	5	25.0	353	14	092269	NIB PROTEIN/COAT PROTE	5.80e+02	478	5	25.0	403	2	055795	HYPOTHETICAL 46.4 KD P	5.80e+02
406	5	25.0	354	2	044407	HYPOTHETICAL 39.4 KD P	5.80e+02	479	5	25.0	405	2	095477	PYRIDOXAL-DEPENDENT HI	5.80e+02
407	5	25.0	357	1	073958	PRIMASE SMALL SUBUNIT	5.80e+02	480	5	25.0	405	5	023050	SIMILARITY TO CALPONIN	5.80e+02
408	5	25.0	357	2	085768	HYPOTHETICAL 38.2 KD P	5.80e+02	481	5	25.0	407	2	054149	3-OXOXACYL-[ACYL-CARRIE	5.80e+02
409	5	25.0	358	2	087162	ORF22-23 PROTEIN.	5.80e+02	482	5	25.0	410	5	016988	FIBROIN-4 (FRAGMENT).	5.80e+02
410	5	25.0	360	10	081388	JAB1 PROTEIN.	5.80e+02	483	5	25.0	413	14	083329	CAPSID PROTEIN.	5.80e+02
411	5	25.0	361	2	045521	PREPRO APRM PRECURSOR.	5.80e+02	484	5	25.0	415	2	096354	HYPOTHETICAL 45.8 KD P	5.80e+02
412	5	25.0	361	10	043383	2A6 MRNA.	5.80e+02	485	5	25.0	416	13	042151	DEVELOPMENTAL ORPHAN R	5.80e+02
413	5	25.0	361	2	083294	THERMOSTABLE ALKALINE	5.80e+02	486	5	25.0	418	2	076526	O418 WAS O397.	5.80e+02
414	5	25.0	363	14	086989	THYMIDINE KINASE.	5.80e+02	487	5	25.0	418	13	098863	EVX-2 GENE.	5.80e+02
415	5	25.0	363	5	022326	SIMILAR TOCOLLAGEN.	5.80e+02	488	5	25.0	419	2	008426	GYRASE BETA SUBUNIT (E	5.80e+02
416	5	25.0	364	14	092546	POLYPROTEIN (FRAGMENT)	5.80e+02	489	5	25.0	419	5	077287	SU(P) PROTEIN.	5.80e+02
417	5	25.0	364	14	081597	NS5 (FRAGMENT).	5.80e+02	490	5	25.0	419	2	008403	GYRASE BETA SUBUNIT (E	5.80e+02
418	5	25.0	364	14	092543	POLYPROTEIN (FRAGMENT)	5.80e+02	491	5	25.0	419	1	094949	HYPOTHETICAL 46.7 KD P	5.80e+02
419	5	25.0	364	14	092544	POLYPROTEIN (FRAGMENT)	5.80e+02	492	5	25.0	419	2	031225	GYRASE B SUBUNIT (FRAG	5.80e+02
420	5	25.0	364	14	092537	POLYPROTEIN (FRAGMENT)	5.80e+02	493	5	25.0	420	3	078892	FISSION YEAST (FRAGEN	5.80e+02
421	5	25.0	364	14	092542	POLYPROTEIN (FRAGMENT)	5.80e+02	494	5	25.0	420	2	033464	GYRASE B SUBUNIT (FRAG	5.80e+02
422	5	25.0	364	14	092541	POLYPROTEIN (FRAGMENT)	5.80e+02	495	5	25.0	420	2	033444	GYRASE B SUBUNIT (FRAG	5.80e+02
423	5	25.0	364	14	092539	POLYPROTEIN (FRAGMENT)	5.80e+02	496	5	25.0	420	3	078765	FISSION YEAST (FRAGEN	5.80e+02
424	5	25.0	364	14	092540	POLYPROTEIN (FRAGMENT)	5.80e+02	497	5	25.0	421	5	093119	ANTHERAEA PERNY FIBRO	5.80e+02
425	5	25.0	364	14	092545	POLYPROTEIN (FRAGMENT)	5.80e+02	498	5	25.0	421	2	059713	SALICYLATE HYDROXYLASE	5.80e+02
426	5	25.0	364	14	092538	POLYPROTEIN (FRAGMENT)	5.80e+02	499	5	25.0	422	2	033430	GYRASE B SUBUNIT (FRAG	5.80e+02
427	5	25.0	364	14	039648	RNA FOR NS5, PARTIAL C	5.80e+02	500	5	25.0	422	2	033442	GYRASE B SUBUNIT (FRAG	5.80e+02
428	5	25.0	365	10	042917	1-AMINOCYCLOPROPANE-CA	5.80e+02	501	5	25.0	422	10	038948	5'-ADENYLPHOSPHOSULF	5.80e+02
429	5	25.0	365	14	089196	INTERMEDIATE FILAMENT	5.80e+02	502	5	25.0	422	2	032856	GYRASE B SUBUNIT (FRAG	5.80e+02
430	5	25.0	367	14	089736	POLYPROTEIN (FRAGMENT)	5.80e+02	503	5	25.0	423	14	085456	AVIAN PROVIRAL RETROVI	5.80e+02
431	5	25.0	367	14	089735	POLYPROTEIN (FRAGMENT)	5.80e+02	504	5	25.0	423	2	033452	GYRASE B SUBUNIT (FRAG	5.80e+02
432	5	25.0	367	5	017642	SIMILAR TO MAP KINASES	5.80e+02	505	5	25.0	423	2	033429	GYRASE B SUBUNIT (FRAG	5.80e+02
433	5	25.0	367	3	059847	HAPB.	5.80e+02	506	5	25.0	423	3	014456	CHITINASE CHIT42 (EC 3	5.80e+02
434	5	25.0	369	9	048455	COMPLETE NUCLEOTIDE SE	5.80e+02	507	5	25.0	423	2	033504	GYRASE B SUBUNIT (FRAG	5.80e+02
435	5	25.0	369	5	046009	ZK228.3 PROTEIN.	5.80e+02	508	5	25.0	425	4	075038	KIAA0450 PROTEIN.	5.80e+02
436	5	25.0	370	2	050895	CAROTENOID GENE CLUSTE	5.80e+02	509	5	25.0	425	6	002661	SCO-SPONDIN (FRAGMENT)	5.80e+02
437	5	25.0	371	5	018870	SIMILAR TO THE ATIC FR	5.80e+02	510	5	25.0	426	13	090968	RETINOIC ACID RECEPT	5.80e+02
438	5	25.0	373	1	064635	GALACTOSYL-TRANSFERASE	5.80e+02	511	5	25.0	426	2	068135	HYPOTHETICAL 45.5 KD P	5.80e+02
439	5	25.0	374	10	093907	BZIP TRANSCRIPTION FAC	5.80e+02	512	5	25.0	429	10	023081	SIMILAR TO RECEPTOR KI	5.80e+02
440	5	25.0	374	5	070664	FL1F9.1 PROTEIN.	5.80e+02	513	5	25.0	431	10	023017	TIGL1.7 PROTEIN.	5.80e+02
441	5	25.0	376	14	035259	THYMIDINE KINASE.	5.80e+02	514	5	25.0	433	14	085207	N PROTEIN.	5.80e+02
442	5	25.0	376	10	096565	CARPEIC ACID O-METHYL	5.80e+02	515	5	25.0	434	1	030067	HYPOTHETICAL 50.6 KD P	5.80e+02
443	5	25.0	376	14	072346	THYMIDINE KINASE	5.80e+02	516	5	25.0	434	10	064886	PUTATIVE HEME A-FARNES	5.80e+02
444	5	25.0	376	4	043664	SIMILAR TO P31245.	5.80e+02	517	5	25.0	436	5	046174	HEPATOCYTE NUCLEAR FAC	5.80e+02
445	5	25.0	377	11	084693	EDA PROTEIN HOMOLOG.	5.80e+02	518	5	25.0	438	2	067997	AMTB.	5.80e+02
446	5	25.0	377	13	070057	DELAYED RECIPIER POTA	5.80e+02	519	5	25.0	438	2	049110	PYRUVATE DEHYDROGENASE	5.80e+02
447	5	25.0	378	10	041557	HBP-1-ASSOCIATED LEUCI	5.80e+02	520	5	25.0	438	2	060556	HYPOTHETICAL 48.0 KD P	5.80e+02
448	5	25.0	378	2	041462	PHOTOINDUCIBLE GENE LI	5.80e+02	521	5	25.0	439	11	070623	TRANSCRIPTION FACTOR 1	5.80e+02
449	5	25.0	378	5	016646	F56D12.5 PROTEIN.	5.80e+02	522	5	25.0	440	10	081642	IAA-ALA HYDROLASE.	5.80e+02
450	5	25.0	379	2	089923	PUTATIVE TRANSPPOSASE.	5.80e+02	523	5	25.0	440	2	066119	DIHYDROLIPOAMIDE S-ACE	5.80e+02
451	5	25.0	380	11	088321	ANTISECRETORY FACTOR.	5.80e+02	524	5	25.0	441	10	023053	SIMILAR TO SYNECHOCYST	5.80e+02
452	5	25.0	381	5	019458	F40G12.11 PROTEIN.	5.80e+02	525	5	25.0	442	13	079878	RETINOIC ACID RECEPT	5.80e+02
453	5	25.0	381	2	035392	E1-ALPHA BRANCHED-CHAI	5.80e+02	526	5	25.0	444	10	004373	JR3 PROTEIN.	5.80e+02
454	5	25.0	381	4	075543	HYPOTHETICAL 41.9 KD P	5.80e+02	527	5	25.0	445	5	017204	CO1B2.4 PROTEIN.	5.80e+02
455	5	25.0	383	1	058062	383AA LONG HYPOTHETICA	5.80e+02	528	5	25.0	447	2	074824	HYPOTHETICAL 48.2 KD P	5.80e+02
456	5	25.0	384	5	027500	F49E11.7 PROTEIN.	5.80e+02	529	5	25.0	448	2	071556	HYPOTHETICAL 46.5 KD P	5.80e+02
457	5	25.0	384	1	039044	HYPOTHETICAL PROTEIN M	5.80e+02	530	5	25.0	449	13	092170	BAND 17.	5.80e+02
458	5	25.0	385	2	005881	HYPOTHETICAL 39.0 KD P	5.80e+02	531	5	25.0	454	8	046971	RIBULOSE BISPHOSPHATE	5.80e+02

532	5	25.0	454	4	Q99540	LIGASE-LIKE PROTEIN.	5.80e+02	605	5	25.0	506	2	083983	GLU-TRNA AMIDOTRANSFER	5.80e+02
533	5	25.0	454	8	O46970	RIBULOSE-1,5-BISPHOSPH	5.80e+02	606	5	25.0	506	14	082004	2ND ATG.	5.80e+02
534	5	25.0	455	8	O46972	RIBULOSE BISPHOSPHATE	5.80e+02	607	5	25.0	508	10	Q89951	PROTEIN PHOSPHATASE PP	5.80e+02
535	5	25.0	456	2	O84252	REPLICATION INITIATION	5.80e+02	608	5	25.0	509	2	Q85163	MULTIDRUG RESISTANCE E	5.80e+02
536	5	25.0	457	4	Q13518	PROTEIN KINASE C-BINDI	5.80e+02	609	5	25.0	509	11	O08749	DIHYDROLIPOAMIDE DEHYD	5.80e+02
537	5	25.0	458	2	P71354	TRK SYSTEM POTASSIUM U	5.80e+02	610	5	25.0	509	2	O05457	HYPOTHETICAL 53.2 KD P	5.80e+02
538	5	25.0	459	10	Q35584	DNA-BINDING PROTEIN HO	5.80e+02	611	5	25.0	510	10	O40938	SUCROSE TRANSPORTER.	5.80e+02
539	5	25.0	459	8	Q37266	RIBULOSE BISPHOSPHATE	5.80e+02	612	5	25.0	511	14	O91332	NUCLEAR ANTIGEN ERNA-1	5.80e+02
540	5	25.0	459	10	O82679	ACC SYNTHASE (EC 4.4.1	5.80e+02	613	5	25.0	511	1	O29309	GLUTAMATE SYNTHASE (GL	5.80e+02
541	5	25.0	462	5	Q09958	HYPOTHETICAL 52.1 KD P	5.80e+02	614	5	25.0	511	4	Q14131	DIHYDROLIPOAMIDE DEHYD	5.80e+02
542	5	25.0	463	2	O53416	PGRS-FAMILY PROTEIN.	5.80e+02	615	5	25.0	512	2	Q33238	HYPOTHETICAL 55.1 KD P	5.80e+02
543	5	25.0	465	3	O00092	PHYTASE PRECURSOR (EC	5.80e+02	616	5	25.0	512	2	O08251	LIPOPROTEIN.	5.80e+02
544	5	25.0	465	13	O42270	P52 PRO-APOTOTIC PROTE	5.80e+02	617	5	25.0	514	5	O26342	HISTIDYL TRNA SYNTHETA	5.80e+02
545	5	25.0	465	2	O32890	HYPOTHETICAL 49.8 KD P	5.80e+02	618	5	25.0	514	2	O54426	TYRR.	5.80e+02
546	5	25.0	466	3	O00096	PHYTASE (EC 3.1.3.8) (	5.80e+02	619	5	25.0	515	13	O12941	TOM-1B PROTEIN.	5.80e+02
547	5	25.0	466	2	O50645	GLUTAMATE DECARBOXYLAS	5.80e+02	620	5	25.0	515	2	P96182	FLAVOXYTOCHROME C FLAV	5.80e+02
548	5	25.0	467	4	Q15626	STEROID RECEPTOR (TR2-	5.80e+02	621	5	25.0	518	10	O81301	T14P8.1 PROTEIN.	5.80e+02
549	5	25.0	467	2	P73043	GLUTAMATE DECARBOXYLAS	5.80e+02	622	5	25.0	519	2	O44372	OPHA.	5.80e+02
550	5	25.0	468	8	Q32863	RIBULOSE BISPHOSPHATE	5.80e+02	623	5	25.0	519	14	O41671	POLYPROTEIN (FRAGMENTS	5.80e+02
551	5	25.0	470	2	P95121	HYPOTHETICAL 46.7 KD P	5.80e+02	624	5	25.0	519	14	O41672	POLYPROTEIN (FRAGMENTS	5.80e+02
552	5	25.0	470	2	O31101	OUTER MEMBRANE CHANNEL	5.80e+02	625	5	25.0	519	14	O41678	POLYPROTEIN (FRAGMENTS	5.80e+02
553	5	25.0	471	4	Q99518	FLAVIN-CONTAINING MONO	5.80e+02	626	5	25.0	519	14	O41679	POLYPROTEIN (FRAGMENTS	5.80e+02
554	5	25.0	472	11	Q61066	ADRENAL HYPOPLASIA, CO	5.80e+02	627	5	25.0	519	14	O41677	POLYPROTEIN (FRAGMENTS	5.80e+02
555	5	25.0	473	8	Q36770	RIBULOSE BISPHOSPHATE	5.80e+02	628	5	25.0	519	14	O41675	POLYPROTEIN (FRAGMENTS	5.80e+02
556	5	25.0	473	5	O4072	ASTACUS EGG ASTACIN PR	5.80e+02	629	5	25.0	519	14	O41674	POLYPROTEIN (FRAGMENTS	5.80e+02
557	5	25.0	473	8	Q36811	RIBULOSE BISPHOSPHATE	5.80e+02	630	5	25.0	519	14	O41668	POLYPROTEIN (FRAGMENTS	5.80e+02
558	5	25.0	474	10	O04930	UDP-GLUCOSE:INDOLE-3-A	5.80e+02	631	5	25.0	519	14	O41669	POLYPROTEIN (FRAGMENTS	5.80e+02
559	5	25.0	474	2	O67007	NADH OXIDASE.	5.80e+02	632	5	25.0	519	14	O41676	POLYPROTEIN (FRAGMENTS	5.80e+02
560	5	25.0	474	10	O23282	PUTATIVE FLAVONOL 3-O-	5.80e+02	633	5	25.0	519	14	O41670	POLYPROTEIN (FRAGMENTS	5.80e+02
561	5	25.0	474	8	O46976	RIBULOSE-1,5-BISPHOSPH	5.80e+02	634	5	25.0	520	2	O48440	4-HYDROXYPHENYLACETATE	5.80e+02
562	5	25.0	476	2	O56790	HRPXV.	5.80e+02	635	5	25.0	520	5	O22408	T11G6.1 PROTEIN.	5.80e+02
563	5	25.0	476	2	O82880	HRPXT PROTEIN.	5.80e+02	636	5	25.0	524	10	O65504	HYPOTHETICAL 58.2 KD P	5.80e+02
564	5	25.0	476	2	O69097	REGULATORY PROTEIN HRP	5.80e+02	637	5	25.0	525	2	Q52019	KETOGUTARATE SEMIALDE	5.80e+02
565	5	25.0	477	1	O26442	SUCCINOGLYCAN BIOSYNTH	5.80e+02	638	5	25.0	526	5	O24044	FIZZY (FZY).	5.80e+02
566	5	25.0	477	2	O45516	ALPHA-AMYLASE.	5.80e+02	639	5	25.0	527	5	O02194	PRESENILIN.	5.80e+02
567	5	25.0	477	2	O80676	SUCROSE HYDROLASE (EC	5.80e+02	640	5	25.0	527	5	O76802	PRESENILIN SHORT FORM.	5.80e+02
568	5	25.0	477	2	O45520	2633 ALPHA-AMYLASE.	5.80e+02	641	5	25.0	528	5	O60978	L3162.7	5.80e+02
569	5	25.0	479	5	P91625	LADYBIRD EARLY HOMEODO	5.80e+02	642	5	25.0	529	13	Q98870	MYOCYTE ENHANCER FACTO	5.80e+02
570	5	25.0	480	2	Q30081	HYPOTHETICAL 51.4 KD P	5.80e+02	643	5	25.0	529	5	O62254	F49A5.8 PROTEIN.	5.80e+02
571	5	25.0	481	10	O81498	F9D12.4 PROTEIN.	5.80e+02	644	5	25.0	529	1	Q48937	F4ME, FMDF, FMDA, FMDC	5.80e+02
572	5	25.0	481	2	P71915	HYPOTHETICAL 52.3 KD P	5.80e+02	645	5	25.0	531	10	O40967	BASIC LEUCINE ZIPPER P	5.80e+02
573	5	25.0	481	10	O82381	T27A16.15 PROTEIN.	5.80e+02	646	5	25.0	532	10	O64851	T1D16.17 PROTEIN.	5.80e+02
574	5	25.0	481	2	O66064	ENDOGLUCANASE.	5.80e+02	647	5	25.0	532	10	O22416	TUBULIN UN13.	5.80e+02
575	5	25.0	482	10	Q43810	1-AMINOCYCLOPROPANE 1-	5.80e+02	648	5	25.0	534	3	O13638	MEIOSIS INDUCTION PROT	5.80e+02
576	5	25.0	483	5	O20696	COSMID F53BL.	5.80e+02	649	5	25.0	534	14	O82003	1ST ATG.	5.80e+02
577	5	25.0	483	2	Q56832	HRPXO.	5.80e+02	650	5	25.0	535	2	O06831	HYPOTHETICAL 58.1 KD P	5.80e+02
578	5	25.0	484	2	O59095	BENZALDEHYDE DEHYDROGE	5.80e+02	651	5	25.0	537	5	O27005	54-KDA ANTIGEN (FRAGME	5.80e+02
579	5	25.0	484	2	O56813	RESPONSIBLE FOR XANTHA	5.80e+02	652	5	25.0	537	2	O84781	ACYLGLYCEROPHOSPHOETHA	5.80e+02
580	5	25.0	485	2	O67746	HYPOTHETICAL 56.9 KD P	5.80e+02	653	5	25.0	538	2	O69379	ACTA (FRAGMENT).	5.80e+02
581	5	25.0	485	2	P95165	COSMID Y359.	5.80e+02	654	5	25.0	541	3	O59738	MAJOR FACILITATOR SUPE	5.80e+02
582	5	25.0	485	11	O88998	PANCORTIN-1.	5.80e+02	655	5	25.0	541	10	O65305	ACETOHYDROXYACID SYNTH	5.80e+02
583	5	25.0	486	2	O32911	6-PHOSPHOGLUCONATE DEH	5.80e+02	656	5	25.0	541	5	O02395	PRESENILIN HOMOLOG.	5.80e+02
584	5	25.0	488	1	O29505	HYPOTHETICAL 56.5 KD P	5.80e+02	657	5	25.0	542	5	O77059	BLU-LIGHT RECEPTOR.	5.80e+02
585	5	25.0	489	2	P72455	NADH-N-AMIDINO-SCYLLO-	5.80e+02	658	5	25.0	544	2	O05598	PKS16 PROTEIN.	5.80e+02
586	5	25.0	489	10	O65209	1-AMINOCYCLOPROPANE-1-	5.80e+02	659	5	25.0	545	5	O16547	C35A11.4 PROTEIN.	5.80e+02
587	5	25.0	490	2	P74027	HYPOTHETICAL 52.7 KD P	5.80e+02	660	5	25.0	546	11	P70182	PHOSPHATIDYLINOSITOL 4	5.80e+02
588	5	25.0	490	2	O26836	VARIANT SURFACE GLYCOP	5.80e+02	661	5	25.0	550	2	O05458	HYPOTHETICAL 55.6 KD P	5.80e+02
589	5	25.0	490	10	Q43813	1-AMINOCYCLOPROPANE-1-	5.80e+02	662	5	25.0	557	11	Q63045	PUT. ANTI-MULLERIAN HO	5.80e+02
590	5	25.0	490	2	O84709	GTPASE/GTP-BINDING PRO	5.80e+02	663	5	25.0	558	2	O53912	A201A-RESISTANCE ATP-B	5.80e+02
591	5	25.0	491	2	P96557	PCZA361.26.	5.80e+02	664	5	25.0	559	2	O51513	PHA-SYNTHASE1	5.80e+02
592	5	25.0	491	5	O02035	ECDSONE RECEPTOR.	5.80e+02	665	5	25.0	559	3	O43077	SERINE/THREONINE PROTE	5.80e+02
593	5	25.0	493	10	O82123	ACC SYNTHASE.	5.80e+02	666	5	25.0	561	5	O27007	ROP 2.	5.80e+02
594	5	25.0	493	10	O24668	1-AMINOCYCLOPROPANE-1-	5.80e+02	667	5	25.0	561	13	Q90711	AE2-2 ANION EXCHANGER	5.80e+02
595	5	25.0	493	3	P79077	DELTA-9 FATTY ACID DES	5.80e+02	668	5	25.0	562	14	O10709	GENOMIC RNA, PARTIAL C	5.80e+02
596	5	25.0	496	10	O81102	GLUTAMATE DECARBOXYLAS	5.80e+02	669	5	25.0	562	14	O10723	GENOMIC RNA, PARTIAL C	5.80e+02
597	5	25.0	496	10	O81101	GLUTAMATE DECARBOXYLAS	5.80e+02	670	5	25.0	562	14	O10724	GENOMIC RNA, PARTIAL C	5.80e+02
598	5	25.0	498	10	Q39031	PROTEIN KINASE.	5.80e+02	671	5	25.0	562	14	O10703	GENOMIC RNA, PARTIAL C	5.80e+02
599	5	25.0	498	10	O22335	ACC SYNTHASE (EC 4.4.1	5.80e+02	672	5	25.0	562	14	O10697	GENOMIC RNA, PARTIAL C	5.80e+02
600	5	25.0	498	10	O22334	ACC SYNTHASE (EC 4.4.1	5.80e+02	673	5	25.0	562	14	O10717	GENOMIC RNA, PARTIAL C	5.80e+02
601	5	25.0	499	13	Q91392	RETINOIC ACID RECEPTOR	5.80e+02	674	5	25.0	567	1	O75082	MTGB-RELATED PROTEIN M	5.80e+02
602	5	25.0	502	2	O52558	VSRA.	5.80e+02	675	5	25.0	567	4	O26602	HYPOTHETICAL 65.3 KD P	5.80e+02
603	5	25.0	503	2	Q58801	HRPXC DNA (FRAGMENT).	5.80e+02	676	5	25.0	571	5	O21266	SIMILAR TO THE TYPE-B	5.80e+02
604	5	25.0	505	5	O04651	C06B3.3 PROTEIN.	5.80e+02	677	5	25.0	574	10	O22305	PEPTIDE TRANSPORTER.	5.80e+02

678	5	25.0	575	5	Q27604	SGG39 PROTEIN KINASE (CYCD.	5.80e+02	751	5	25.0	660	13	073925	KVLQT POTASSIUM CHANNE	5.80e+02
679	5	25.0	576	2	O06137		5.80e+02	752	5	25.0	661	11	070205	DREBRIN E (FRAGMENT).	5.80e+02
680	5	25.0	579	2	O33057	HYPOTHETICAL 61.7 KD P	5.80e+02	753	5	25.0	663	5	077001	HELICASE PITCHOUNE.	5.80e+02
681	5	25.0	580	5	P09696	5'-NUCLEOTIDASE (FRAGM	5.80e+02	754	5	25.0	668	14	008177	COAT PROTEIN (CAPSID P	5.80e+02
682	5	25.0	581	10	Q43111	PECTINSTERASE (EC 3.1	5.80e+02	755	5	25.0	673	2	051776	EXCINUCLEASE ABC. SUBU	5.80e+02
683	5	25.0	581	2	O34023	HYPOTHETICAL 66.2 KD P	5.80e+02	756	5	25.0	674	4	043243	ADCUINUCLEASE ABC. SUBU	5.80e+02
684	5	25.0	587	5	O77154	NUCLEAR RECEPTOR 1.	5.80e+02	757	5	25.0	675	14	089273	MC106L.	5.80e+02
685	5	25.0	590	11	O60927	TR2 (MTR2).	5.80e+02	758	5	25.0	675	4	092773	ADDUCIN GAMMA SUBUNIT.	5.80e+02
686	5	25.0	590	11	P97763	ORPHAN RECEPTOR.	5.80e+02	759	5	25.0	675	2	O57071	PTS SYSTEM, GLUCOSE-SP	5.80e+02
687	5	25.0	590	11	O62152	ORPHAN RECEPTOR, TR2-1	5.80e+02	760	5	25.0	676	4	000347	KIDNEY AND CARDIAC VOL	5.80e+02
688	5	25.0	591	2	O69736	HYPOTHETICAL 64.6 KD P	5.80e+02	761	5	25.0	676	2	047759	LOW AFFINITY PENICILLIN	5.80e+02
689	5	25.0	592	2	O69316	(FRAGMENT).	5.80e+02	762	5	25.0	679	1	059078	679AA LONG HYPOTHETICA	5.80e+02
690	5	25.0	592	3	O60008	60 KD CHAPERONIN (PROT	5.80e+02	763	5	25.0	686	4	015195	VILLIN-LIKE PROTEIN.	5.80e+02
691	5	25.0	592	4	O13856	5-KINOIDAMAZOLE-4-CAR	5.80e+02	764	5	25.0	687	11	089024	CCYC PROTEIN.	5.80e+02
692	5	25.0	592	2	O72311	TRANSCRIPTIONAL ACTIVA	5.80e+02	765	5	25.0	688	11	088934	RF-C/ACTIVATOR 1 HOMOL	5.80e+02
693	5	25.0	592	2	O69184	PUTATIVE TRANSCRIPTION	5.80e+02	766	5	25.0	692	2	O53922	PTS SYSTEM, GLUCOSE-SP	5.80e+02
694	5	25.0	592	11	O35567	5-AMINOIMIDAZOLE-4-CAR	5.80e+02	767	5	25.0	693	4	032607	MYELOBLAST KIAA0208.	5.80e+02
695	5	25.0	596	2	O60029	BETA-1,1,4-EXOCCELLULASE	5.80e+02	768	5	25.0	696	14	002684	A-TYPE INCLUSION BODY	5.80e+02
696	5	25.0	597	11	O63615	VACUOLAR PROTEIN SORTI	5.80e+02	769	5	25.0	696	14	002685	A-TYPE INCLUSION BODY	5.80e+02
697	5	25.0	600	2	P66890	BCCA.	5.80e+02	770	5	25.0	697	5	061121	RNA GUANYLYLTRANSFERAS	5.80e+02
698	5	25.0	601	2	O83801	FUMARATE REDUCTASE FLA	5.80e+02	771	5	25.0	699	2	O07219	HYPOTHETICAL 72.4 KD P	5.80e+02
699	5	25.0	603	4	O15625	STEROID RECEPTOR (TR2-	5.80e+02	772	5	25.0	700	8	033181	NADH DEHYDROGENASE SUB	5.80e+02
700	5	25.0	604	10	O24404	GLUCOSE-6-PHOSPHATE 1-	5.80e+02	773	5	25.0	701	14	065588	HYPOTHETICAL 72.6 KD P	5.80e+02
701	5	25.0	605	2	O32214	YVGR PROTEIN.	5.80e+02	774	5	25.0	702	14	085393	HOMOLOG OF VACCINIA WR	5.80e+02
702	5	25.0	606	5	O72215	NUCLEAR HORMONE RECEPT	5.80e+02	775	5	25.0	702	14	089171	ORF2L.	5.80e+02
703	5	25.0	608	2	O87017	ALPHA-GLYCEROPHOSPHATE	5.80e+02	776	5	25.0	707	2	069998	PUTATIVE DNA GYRASE SU	5.80e+02
704	5	25.0	610	2	P73400	ABC TRANSPORTER.	5.80e+02	777	5	25.0	708	2	084092	LOW CALCIUM RESPONSE D	5.80e+02
705	5	25.0	610	10	O39828	SDLSA.	5.80e+02	778	5	25.0	712	14	037167	POLYPROTEIN (FRAGMENT)	5.80e+02
706	5	25.0	611	2	O69514	DNA POLYMERASE III SUB	5.80e+02	779	5	25.0	712	14	037168	POLYPROTEIN (FRAGMENT)	5.80e+02
707	5	25.0	614	2	O84147	SERINE/THREONINE PROTE	5.80e+02	780	5	25.0	712	3	Q12643	ACETYLGUTAMATE SYNTHA	5.80e+02
708	5	25.0	615	13	O93575	REELIN (FRAGMENT).	5.80e+02	781	5	25.0	715	2	P74517	LIPOPROTEIN NLDP.	5.80e+02
709	5	25.0	615	10	O82002	PUTATIVE CULLIN PROTEI	5.80e+02	782	5	25.0	716	4	Q13531	DISHVELLED (DVL).	5.80e+02
710	5	25.0	615	9	O21976	GP49.	5.80e+02	783	5	25.0	716	11	061062	DISHVELLED 3.	5.80e+02
711	5	25.0	617	2	O33848	FERREDOXIN OXIDOREDUCT	5.80e+02	784	5	25.0	716	4	Q14642	DISHVELLED 3.	5.80e+02
712	5	25.0	618	1	O27880	INDOLEPYRUVATE OXIDORE	5.80e+02	785	5	25.0	716	4	Q14164	KIAA0151 PROTEIN.	5.80e+02
713	5	25.0	620	5	O16519	TOSH4.14 PROTEIN.	5.80e+02	786	5	25.0	716	4	Q92997	DISHVELLED 3 (DVL3).	5.80e+02
714	5	25.0	621	10	O49047	FLAVANONE 3-HYDROXYLAS	5.80e+02	787	5	25.0	724	11	O62428	TRANSPORTER 1, ABC (AT	5.80e+02
715	5	25.0	623	2	P95072	HYPOTHETICAL 65.9 KD P	5.80e+02	788	5	25.0	724	11	O62429	TRANSPORTER 1, ABC (AT	5.80e+02
716	5	25.0	623	5	O17562	C03E10.5 PROTEIN.	5.80e+02	789	5	25.0	725	11	P97949	TAP1 PROTEIN.	5.80e+02
717	5	25.0	624	2	Q51708	POLY(3-HYDROXYALKANOAT	5.80e+02	790	5	25.0	725	11	P97560	TAP1 PROTEIN.	5.80e+02
718	5	25.0	625	4	Q44677	KIAA0171 PROTEIN.	5.80e+02	791	5	25.0	725	11	P97559	TAP1 PROTEIN.	5.80e+02
719	5	25.0	627	5	O15832	L3169.3.	5.80e+02	792	5	25.0	727	10	P93755	CYS3HS ZINC FINGER PR	5.80e+02
720	5	25.0	627	2	O86945	SUB-UNIT I OF AN AA3-T	5.80e+02	793	5	25.0	728	2	P96281	HYPOTHETICAL 75.3 KD P	5.80e+02
721	5	25.0	627	1	O30244	HYPOTHETICAL 72.0 KD P	5.80e+02	794	5	25.0	728	9	O38369	70 KD STRUCTURAL PROTE	5.80e+02
722	5	25.0	627	2	P97041	DNA FOR 5S RNA.	5.80e+02	795	5	25.0	730	5	O61089	HYPOTHETICAL 80.0 KD P	5.80e+02
723	5	25.0	634	14	Q06122	CYTOPLASMIC INCLUSION	5.80e+02	796	5	25.0	730	14	Q06634	TRANS-ACTING TRANSCRIP	5.80e+02
724	5	25.0	634	4	O45209	BING1.	5.80e+02	797	5	25.0	732	10	Q04289	PHOSPHATIDYLINOSITOL 3	5.80e+02
725	5	25.0	634	5	O62799	AI.12/9 ANTIGEN (FRAGM	5.80e+02	798	5	25.0	737	6	Q28956	17BETA-ESTRADIOL DEHYD	5.80e+02
726	5	25.0	635	2	O65065	CMC-XYLANASE (FRAGMENT	5.80e+02	799	5	25.0	741	2	O56641	HAEMOLYSIN.	5.80e+02
727	5	25.0	636	14	O69104	PROTEASE (FRAGMENT).	5.80e+02	800	5	25.0	741	6	Q28434	ABC-TRANSPORTER (FRAGM	5.80e+02
728	5	25.0	636	10	O81892	PUTATIVE ASPARTATE--TR	5.80e+02	801	5	25.0	747	6	Q28432	ABC-TRANSPORTER (FRAGM	5.80e+02
729	5	25.0	636	5	Q15987	FIBROIN-3 (FRAGMENT).	5.80e+02	802	5	25.0	748	6	Q28433	ABC-TRANSPORTER.	5.80e+02
730	5	25.0	637	14	Q070814	POLYPROTEIN (FRAGMENT).	5.80e+02	803	5	25.0	748	6	Q65820	UL46.	5.80e+02
731	5	25.0	638	14	O61730	OCTOPAMINE RECEPTOR OA	5.80e+02	804	5	25.0	748	4	Q16149	PEPTIDE TRANSPORTER (T	5.80e+02
732	5	25.0	638	14	O95927	UT26 PROTEASE.	5.80e+02	805	5	25.0	751	10	O50072	HYPOTHETICAL 86.9 KD P	5.80e+02
733	5	25.0	638	10	O41769	ACETOHYDROXYACID SYNTH	5.80e+02	806	5	25.0	751	5	O60975	L3162-2.	5.80e+02
734	5	25.0	638	10	O41768	ACETOHYDROXYACID SYNTH	5.80e+02	807	5	25.0	755	4	O14715	RAN BINDING PROTEIN 2	5.80e+02
735	5	25.0	642	2	Q55103	CHO-ORF2.	5.80e+02	808	5	25.0	757	2	O52177	(P)PPGPP SYNTHETASE.	5.80e+02
736	5	25.0	644	11	Q63793	PC7A (FRAGMENT).	5.80e+02	809	5	25.0	760	10	O80856	PUTATIVE SWHL PROTEIN.	5.80e+02
737	5	25.0	645	4	Q15673	PLECKSTRIN-SEC7 DOMAIN	5.80e+02	810	5	25.0	763	3	Q12653	BETA-GLUCOSIDASE (EC 3	5.80e+02
738	5	25.0	648	14	Q95714	DNA BINDING PROTEIN.	5.80e+02	811	5	25.0	765	2	O87373	C-DI-GMP PHOSPHODIESTE	5.80e+02
739	5	25.0	648	14	O98826	PROVIRAL DNA, HSRV1 DE	5.80e+02	812	5	25.0	765	2	O54183	CHITINASE.	5.80e+02
740	5	25.0	648	14	P89871	GAG PROTEIN PRECURSOR.	5.80e+02	813	5	25.0	767	13	Q91215	SOX-L2.	5.80e+02
741	5	25.0	648	10	O64406	ENVELOPE-LIKE PROTEIN	5.80e+02	814	5	25.0	767	3	Q13790	PUTATIVE CELL DIVISION	5.80e+02
742	5	25.0	652	2	Q21934	R1LD1.1.	5.80e+02	815	5	25.0	769	2	Q24889	TRANSCRIPTIONAL REGULA	5.80e+02
743	5	25.0	656	2	Q55710	THIG PROTEIN.	5.80e+02	816	5	25.0	769	10	O49978	SUCROSE-PHOSPHATE SYNT	5.80e+02
744	5	25.0	656	2	Q55064	RUVB.	5.80e+02	817	5	25.0	769	5	O61348	BOBBY SOX.	5.80e+02
745	5	25.0	657	10	O22835	SF16 ISOLOG.	5.80e+02	818	5	25.0	776	4	O60719	CULLIN 1.	5.80e+02
746	5	25.0	659	5	O94600	L5701.3.	5.80e+02	819	5	25.0	779	11	O88705	HYPERPOLARIZATION-ACTI	5.80e+02
747	5	25.0	659	2	P72587	ARGININE DECARBOXYLASE	5.80e+02	820	5	25.0	779	5	Q22801	COSMID T26A5	5.80e+02
748	5	25.0	660	14	O81878	COMPLETE GENOME SEQUEN	5.80e+02	821	5	25.0	783	10	O80558	SC2J18.21 PROTEIN.	5.80e+02
749	5	25.0	660	14	O69895	ORF-2.	5.80e+02	822	5	25.0	789	11	O63540	TCAI PROTEIN.	5.80e+02
750	5	25.0	660	14	O81871	ORF 2 PRECURSOR.	5.80e+02	823	5	25.0	789	5	Q17219	F53A3.2 PROTEIN.	5.80e+02

824	5	25.0	796	2	069908	HYPOTHETICAL 87.4 KD P	5.80e+02	897	5	25.0	1185	4	Q13135	LAR-INTERACTING PROTEI	5.80e+02
825	5	25.0	799	3	Q08921	CHROMOSOME XVI READING	5.80e+02	898	5	25.0	1186	5	Q61080	MYOSIN IC HEAVY CHAIN.	5.80e+02
826	5	25.0	800	2	Q08860	CELLULOSE SYNTHASE SUB	5.80e+02	899	5	25.0	1201	5	Q24240	ACTIN BINDING PROTEIN.	5.80e+02
827	5	25.0	802	11	Q08894	TAP1 (FRAGMENT)	5.80e+02	900	5	25.0	1202	4	Q13136	LAR-INTERACTING PROTEI	5.80e+02
828	5	25.0	805	5	Q01576	SIMILAR TO NUCLEOPORIN	5.80e+02	901	5	25.0	1219	13	Q07010	AE2-1 ANION EXCHANGER.	5.80e+02
829	5	25.0	806	2	P95175	HYPOTHETICAL 85.4 KD P	5.80e+02	902	5	25.0	1239	3	Q13394	CLASS IV CHITIN SYNTHA	5.80e+02
830	5	25.0	810	2	Q69650	PONA PROTEIN.	5.80e+02	903	5	25.0	1263	5	P91639	SPALT-RELATED GENE.	5.80e+02
831	5	25.0	819	2	Q31582	YFHO PROTEIN.	5.80e+02	904	5	25.0	1271	4	Q12842	BREAKPOINT CLUSTER REG	5.80e+02
832	5	25.0	824	5	Q17763	F01G10.5 PROTEIN.	5.80e+02	905	5	25.0	1277	4	Q12770	KIAA0199 PROTEIN (FRAG	5.80e+02
833	5	25.0	830	4	Q14162	KIAA0149 PROTEIN.	5.80e+02	906	5	25.0	1293	10	Q04251	PREDICTED PROTEIN OF U	5.80e+02
834	5	25.0	830	4	Q43701	ACEYL LDL RECEPTOR PR	5.80e+02	907	5	25.0	1313	3	P87141	PUTATIVE GUANINE-NUCLE	5.80e+02
835	5	25.0	834	5	P27291	SENSORY TRANSDUCTION H	5.80e+02	908	5	25.0	1329	2	Q06810	HYPOTHETICAL 107.4 KD	5.80e+02
836	5	25.0	835	5	Q22469	T13H5.4 PROTEIN.	5.80e+02	909	5	25.0	1336	14	Q73550	POLYPROTEIN.	5.80e+02
837	5	25.0	843	5	Q61123	19S CAP PROTEASOME S2	5.80e+02	910	5	25.0	1371	2	P73337	SENSORY TRANSDUCTION H	5.80e+02
838	5	25.0	846	2	Q05734	SURFACE PROTEIN LK90.	5.80e+02	911	5	25.0	1393	4	Q75872	RAB3-GAP REGULATORY DO	5.80e+02
839	5	25.0	850	14	Q70003	ENVELOPE GLYCOPROTEIN	5.80e+02	912	5	25.0	1396	2	Q84316	RNA POLYMERASE BETA'. ORF.	5.80e+02
840	5	25.0	850	5	Q21350	K08F4.1 PROTEIN.	5.80e+02	913	5	25.0	1401	4	Q15021	PEPTIDE SYNTHASE.	5.80e+02
841	5	25.0	869	13	Q42126	VITELLOGENIN RECEPTOR.	5.80e+02	914	5	25.0	1407	2	Q59826	PUTATIVE CYTOSKELETON	5.80e+02
842	5	25.0	869	2	Q33141	HYPOTHETICAL 95.6 KD P	5.80e+02	915	5	25.0	1420	3	Q37336	PROTEIN-TYROSINE PHOSP	5.80e+02
843	5	25.0	870	4	Q00308	WMP2.	5.80e+02	916	5	25.0	1422	13	Q98936	TOLLOID RELATED-1. TOLKIN.	5.80e+02
844	5	25.0	871	10	Q64892	POLYPROTEIN (FRAGMENT)	5.80e+02	917	5	25.0	1464	5	Q23995	INSECTICIDAL TOXIN REC	5.80e+02
845	5	25.0	875	14	Q56221	P110.	5.80e+02	918	5	25.0	1464	5	Q24132	F15D4.1 PROTEIN.	5.80e+02
846	5	25.0	886	11	Q54951	SEMAPHORIN N (SEMAPHOR	5.80e+02	919	5	25.0	1528	5	P81137	HYPOTHETICAL 159.8 KD	5.80e+02
847	5	25.0	886	2	Q06768	PUTATIVE ATP /GTP-BIND	5.80e+02	920	5	25.0	1529	5	Q45899	PROLIFERATION POTENTIA	5.80e+02
848	5	25.0	890	10	Q04246	PUTATIVE CURLLEAF-LIK	5.80e+02	921	5	25.0	1539	2	P36271	RETINOBLASTOMA BINDING	5.80e+02
849	5	25.0	900	10	Q49160	P105.	5.80e+02	922	5	25.0	1560	11	P97868	FAMILY 10 XYLANASE (EC	5.80e+02
850	5	25.0	901	2	Q83728	PYRUVATE, PHOSPHATE DI	5.80e+02	923	5	25.0	1587	11	P70287	COSMID C56C10.	5.80e+02
851	5	25.0	902	4	Q43632	SPINDLE POLE BODY PROT	5.80e+02	924	5	25.0	1595	2	Q52373	LCFACAS5.	5.80e+02
852	5	25.0	914	4	Q99493	SCA2 PROTEIN (FRAGMENT	5.80e+02	925	5	25.0	1599	5	Q18892	PUTATIVE CATION-TRANSP	5.80e+02
853	5	25.0	925	5	Q93238	C17G1.4 PROTEIN.	5.80e+02	926	5	25.0	1607	5	Q94599	KIAA0533 PROTEIN (LAMI	5.80e+02
854	5	25.0	932	11	Q62030	PACE4A (FRAGMENT).	5.80e+02	927	5	25.0	1609	2	Q53114	186K PROTEIN.	5.80e+02
855	5	25.0	932	13	Q57587	TENASCIN W PRECURSOR.	5.80e+02	928	5	25.0	1645	4	Q15230	FLN21.18.	5.80e+02
856	5	25.0	939	2	P28179	FORMATE DEHYDROGENASE	5.80e+02	929	5	25.0	1648	14	P89877	MICROTUBULE BINDING PR	5.80e+02
857	5	25.0	943	5	Q02003	NEUROTRANSMITTER TRANS	5.80e+02	930	5	25.0	1690	10	Q49278	HYPOTHETICAL 185.1 KD	5.80e+02
858	5	25.0	947	11	Q08689	CNR1.	5.80e+02	931	5	25.0	1699	5	Q44929	PUTATIVE OUTER MEMBRAN	5.80e+02
859	5	25.0	948	4	Q15290	RB PROTEIN BINDING PRO	5.80e+02	932	5	25.0	1749	2	P73032	POLYKETIDE SYNTHASE.	5.80e+02
860	5	25.0	970	10	Q08790	REVERSE-TRANSCRIPTASE-	5.80e+02	933	5	25.0	1751	2	Q84448	RIFAMYCIN POLYKETIDE S	5.80e+02
861	5	25.0	970	2	Q05931	HYPOTHETICAL 105.7 KD	5.80e+02	934	5	25.0	1762	2	Q52546	T28B8.4 PROTEIN.	5.80e+02
862	5	25.0	977	5	Q44322	REVERSE TRANSCRIPTASE	5.80e+02	935	5	25.0	1763	2	Q52790	MAJOR MEROZOITE SURFAC	5.80e+02
863	5	25.0	978	2	P96104	DIHYDROLIPOYL TRANSACE	5.80e+02	936	5	25.0	1779	5	Q18150	NONSTRUCTURAL POLYPROT	5.80e+02
864	5	25.0	982	2	Q33492	STYRENE SENSOR KINASE.	5.80e+02	937	5	25.0	1785	5	Q25685	MANNUROAN C-5-EPIMERA	5.80e+02
865	5	25.0	988	14	Q87041	STVCPZ, COMPLETE GENOM	5.80e+02	938	5	25.0	1789	14	Q83883	TYLACTONE SYNTHASE MOD	5.80e+02
866	5	25.0	988	14	Q98830	PROVIRAL DNA, HSRV1 DE	5.80e+02	939	5	25.0	1839	2	Q44496	F36F2.3 PROTEIN.	5.80e+02
867	5	25.0	988	4	Q14655	HYPOTHETICAL 109.6 KD	5.80e+02	940	5	25.0	1841	2	Q33958	NON-STRUCTURAL POLYPRO	5.80e+02
868	5	25.0	989	14	P90288	ENVELOPE PROTEIN.	5.80e+02	941	5	25.0	1877	5	Q62235	PROTEIN TYROSINE PHOSP	5.80e+02
869	5	25.0	997	2	Q44495	MANNUROAN C-5-EPIMERA	5.80e+02	942	5	25.0	1880	14	Q93268	NONMUSCLE MYOSIN-II HE	5.80e+02
870	5	25.0	997	2	Q68533	TRANSPONASE.	5.80e+02	943	5	25.0	1948	4	Q13332	COSMID F10G7.	5.80e+02
871	5	25.0	999	4	Q75115	KIAA0618 PROTEIN.	5.80e+02	944	5	25.0	2057	5	Q94987	L3162.1 PROTEIN.	5.80e+02
872	5	25.0	1000	2	Q85086	LACTOFERRIN BINDING PR	5.80e+02	945	5	25.0	2168	5	Q19330	T05A10.1 PROTEIN.	5.80e+02
873	5	25.0	1000	2	Q85082	LACTOFERRIN BINDING PR	5.80e+02	946	5	25.0	2241	5	Q15850	CODED FOR BY C. ELEGAN	5.80e+02
874	5	25.0	1008	5	Q45055	1-EVIDENCE=PREDICTED B	5.80e+02	947	5	25.0	2251	5	Q22190	CODED FOR BY C. ELEGAN	5.80e+02
875	5	25.0	1011	5	Q23978	MYOSIN-1A.	5.80e+02	948	5	25.0	2288	5	Q23081	MLL-AF4 DER(11). FUSTON	5.80e+02
876	5	25.0	1013	2	Q48426	HSDR.	5.80e+02	949	5	25.0	2305	5	P90749	PUTATIVE PHOSPHATIDYLI	5.80e+02
877	5	25.0	1014	5	Q26152	V-SERA 2.	5.80e+02	950	5	25.0	2311	4	Q08524	CERP250 CENTROSOME ASSO	5.80e+02
878	5	25.0	1015	10	Q64673	F22013.25.	5.80e+02	951	5	25.0	2335	3	Q14356	INSECTICIDAL TOXIN COM	5.80e+02
879	5	25.0	1021	5	Q19011	SIMILAR TO DYSTROPHIN.	5.80e+02	952	5	25.0	2442	4	Q05088	FENGICIN SYNTHETASE.	5.80e+02
880	5	25.0	1025	10	Q40640	RECEPTOR KINASE-LIKE P	5.80e+02	953	5	25.0	2442	4	Q14812	VARIANT-SPECIFIC SURFA	5.80e+02
881	5	25.0	1026	10	Q49529	PREDICTED PROTEIN.	5.80e+02	954	5	25.0	2504	2	Q50160	INOSITOL TRIPHOSPHATE	5.80e+02
882	5	25.0	1036	2	Q54483	SSP-H1.	5.80e+02	955	5	25.0	2554	2	Q30981	HUMAN TYPE 3 INOSITOL	5.80e+02
883	5	25.0	1043	5	Q23999	PUTATIVE SINGLE-STRAND	5.80e+02	956	5	25.0	2571	2	Q87704	TYPE 3 INOSITOL 1,4,5-	5.80e+02
884	5	25.0	1046	5	Q22639	SERUM OPACITY FACTOR.	5.80e+02	957	5	25.0	2664	5	Q26033	HUMAN TYPE 1 INOSITOL	5.80e+02
885	5	25.0	1046	5	P91176	CODED FOR BY C. ELEGAN	5.80e+02	958	5	25.0	2670	11	Q63269	POLYPROTEIN.	5.80e+02
886	5	25.0	1047	5	Q61136	CALCIUM MOTIVE P-TYPE	5.80e+02	959	5	25.0	2671	4	Q14573	POLYPROTEIN PRECURSOR.	5.80e+02
887	5	25.0	1052	10	Q23424	DNA CHROMOSOME 4, ESSA	5.80e+02	960	5	25.0	2671	4	Q14649	POLYPROTEIN.	5.80e+02
888	5	25.0	1062	4	Q14688	KIAA0183 PROTEIN (FRAG	5.80e+02	961	5	25.0	2695	4	Q14643	POLYPROTEIN.	5.80e+02
889	5	25.0	1065	5	Q22639	T21C9.2 PROTEIN.	5.80e+02	962	5	25.0	2842	14	Q09804	POLYPROTEIN.	5.80e+02
890	5	25.0	1091	5	Q17550	F54E12.2 PROTEIN.	5.80e+02	963	5	25.0	2842	14	Q09252	POLYPROTEIN.	5.80e+02
891	5	25.0	1096	4	Q60878	DIA-12C PROTEIN.	5.80e+02	964	5	25.0	2842	14	Q36152	POLYPROTEIN.	5.80e+02
892	5	25.0	1101	4	Q60879	DIA-156 PROTEIN.	5.80e+02	965	5	25.0	2842	14	Q36152	POLYPROTEIN.	5.80e+02
893	5	25.0	1115	11	Q89119	PER3 (FRAGMENT).	5.80e+02	966	5	25.0	2842	14	Q36178	POLYPROTEIN.	5.80e+02
894	5	25.0	1121	4	Q35013	KIAA0294.	5.80e+02	967	5	25.0	2842	14	Q36178	POLYPROTEIN.	5.80e+02
895	5	25.0	1122	13	P70047	C2-HC TYPE ZINC FINGER	5.80e+02	968	5	25.0	2842	14	Q36178	POLYPROTEIN.	5.80e+02
896	5	25.0	1170	11	Q61528	ERCC5 PROTEIN.	5.80e+02	969	5	25.0	2842	14	Q36176	POLYPROTEIN.	5.80e+02



970 5 25.0 2873 14 P89967 RNA FOR POLYPROTEIN, C 5.80e+02  
971 5 25.0 2873 14 Q93069 POLYPROTEIN. 5.80e+02  
972 5 25.0 2873 14 Q90481 POLYPROTEIN. 5.80e+02  
973 5 25.0 2873 14 Q93073 POLYPROTEIN. 5.80e+02  
974 5 25.0 2873 14 Q93114 POLYPROTEIN. 5.80e+02  
975 5 25.0 2873 14 Q93074 POLYPROTEIN. 5.80e+02  
976 5 25.0 2873 14 Q93075 POLYPROTEIN. 5.80e+02  
977 5 25.0 2873 14 Q93070 POLYPROTEIN. 5.80e+02  
978 5 25.0 2873 14 Q93071 POLYPROTEIN. 5.80e+02  
979 5 25.0 2873 14 Q93072 POLYPROTEIN. 5.80e+02  
980 5 25.0 2873 14 Q9431 POLYPROTEIN. 5.80e+02  
981 5 25.0 2910 14 Q9433 POLYPROTEIN. 5.80e+02  
982 5 25.0 2933 14 Q9803 POLYPROTEIN. 5.80e+02  
983 5 25.0 2954 14 Q9808 POLYPROTEIN. 5.80e+02  
984 5 25.0 3008 14 Q93929 HCV POLYPROTEIN. 5.80e+02  
985 5 25.0 3023 14 Q94898 GENOME POLYPROTEIN 5.80e+02  
986 5 25.0 3061 14 Q95265 POLYPROTEIN. 5.80e+02  
987 5 25.0 3081 14 Q95105 POLYPROTEIN (FRAGMENT) 5.80e+02  
988 5 25.0 3122 14 P89459 VERY LARGE TEGUMENT PR 5.80e+02  
989 5 25.0 3191 5 Q01335 TOLD3.1 PROTEIN. 5.80e+02  
990 5 25.0 3209 13 Q01374 EXTRACELLULAR REELIN ( 5.80e+02  
991 5 25.0 3229 5 Q26912 PROTEIN 1 OF A DISPERS 5.80e+02  
992 5 25.0 3726 11 Q01329 AT MOTIF-BINDING FACTO 5.80e+02  
993 5 25.0 4005 4 Q13744 ALL-1 PROTEIN. 5.80e+02  
994 5 25.0 4470 14 Q39225 RNA-DIRECTED RNA POLYM 5.80e+02  
995 5 25.0 4578 13 Q42181 PKD1 PROTEIN. 5.80e+02  
996 5 25.0 4845 11 Q98738 UBIQUITIN-CONTUGATING 5.80e+02  
997 5 25.0 5112 5 Q24501 RYANODINE RECEPTOR HOM 5.80e+02  
998 5 25.0 5112 5 Q24500 RYANODINE RECEPTOR HOM 5.80e+02  
999 5 25.0 5126 5 Q24498 RYANODINE RECEPTOR HOM 5.80e+02  
1000 5 25.0 5126 5 Q24499 RYANODINE RECEPTOR HOM 5.80e+02

## ALIGNMENTS

RESULT 1  
ID Q30390 PRELIMINARY; PRT; 351 AA.  
AC Q30390  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE PUTATIVE GONOCOCCAL SENSOR KINASE (FRAGMENT).  
GN PILB.  
OS NEISSERIA MENINGITIDIS.  
OC BACTERIA; PROTEOBACTERIA; BETA SUBDIVISION; NEISSERIACEAE; NEISSERIA.  
OC [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=8013.6;  
RX MEDLINE; 98096361.  
RA ARVIDSON C.G., SO M.;  
RT "Isolation and biochemical characterization of the Pila protein of  
Neisseria meningitidis.";  
RL ARCH. BIOCHEM. BIOPHYS. 348:357-362(1997).  
DR EMBL; AF003940; G2580591; -.  
FT NON\_TER 351  
SQ SEQUENCE 351 AA; 39079 MW; 1F8DF921 CRC32;

Query Match 35.0%; Score 7; DB 2; Length 351;  
Best Local Similarity 100.0%; Pred. No. 7.74e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 203 LPAASG 209  
QY 13 LPAASG 19

RESULT 2  
ID Q23004 PRELIMINARY; PRT; 465 AA.  
AC Q23004  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE \*F6P23.11 PROTEIN.

GN F6P23.11.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA HASEGAWA A.E., SCHUTZ K., DE LA BASTIDE M., LODHI M., GNOJ L.,  
RA GOTTESMAN T., GRANAT S., HAMEED A., KAPLAN N., SHOHDI N.,  
RA VAN KEUREN K., MARRA M., JOHNSON A.F., PARNELL L., DEDHIA N.,  
RA MARTENSEN R., MCCOMBIE W.R.;  
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AC002354; G2264377; -.  
DR PFAM; PF00069; kinase; 1.  
SQ SEQUENCE 465 AA; 52281 MW; 87CA6E06 CRC32;

Query Match 35.0%; Score 7; DB 10; Length 465;  
Best Local Similarity 100.0%; Pred. No. 7.74e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 129 DTLAKRL 135  
QY 7 DTLAKRL 13

RESULT 3  
ID Q42020 PRELIMINARY; PRT; 90 AA.  
AC Q42020  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE PROTEIN KINASE C INHIBITOR HOMOLOGUE (FRAGMENT).  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;  
RA BERTHOUE P., GUERRIER D., GIRAUDAT J.;  
RL SUBMITTED (NOV-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; Z18505; G17628; -.  
DR PFAM; PF00244; 14-3-3; 1.  
DR MENDEL; 6252; Arabid; 1020; 10.  
FT NON\_TER 1  
FT NON\_TER 90  
SQ SEQUENCE 90 AA; 10058 MW; 86C119A6 CRC32;

Query Match 30.0%; Score 6; DB 10; Length 90;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 37 PAAASG 42  
QY 14 PAAASG 19

RESULT 4  
ID P74900 PRELIMINARY; PRT; 99 AA.  
AC P74900  
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE VACUOLAR TYPE ATP SYNTHASE SUBUNIT.  
GN VATL-THERM.  
OS THERMUS AQUATICUS (SUBSP. THERMOPHILUS).  
OC BACTERIA; THERMUS/DEINOCOCCUS GROUP; THERMUS GROUP; THERMUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB8;  
RA YOSHIDA M., YOKOYAMA K.;  
RT \*Operon of Vacuolar type H-ATPase of Thermophilic Eubacterium,



RT Thermus thermophilus."  
RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: D63799; D1010515; -  
SQ SEQUENCE 99 AA; 9836 MW; 14EF409C CRC32;

Query Match 30.0%; Score 6; DB 2; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 AAASGG 28  
|||||  
QY 15 AAASGG 20

RESULT 5  
ID Q26959 PRELIMINARY; PRT; 107 AA.  
AC Q26959;  
DT 01-NOV-1996 (TREMREL. 01, CREATED)  
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE TCPBETA (HJ.3 LOCUS), AGP2BETA-1 AND AGP2BETA-2 GENES.  
GN TCP2BETA.  
OS TRYPAOSOMA CRUZI.  
OC EUKARYOTA; EULENZOIA; KINETOPLASTIDA; TRYPAOSOMATIDAE; TRYPAOSOMA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TULAHUEN;  
RX MEDLINE; 95021509.  
RA VAZQUEZ M.P., SCHIJMAN A.G., LEVIN M.J.;  
RT "A short interspersed repetitive element provides a new 3' acceptor site for trans-splicing in certain ribosomal P2 beta protein genes of Trypanosoma cruzi."  
RL MOL. BIOCHEM. PARASITOL. 64:327-336(1994).  
DR EMBL: X75032; G436146; -  
DR PFAM: PF00428; 60s\_ribosomal; 1.  
SQ SEQUENCE 107 AA; 10411 MW; 00CCE8B4 CRC32;

Query Match 30.0%; Score 6; DB 5; Length 107;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AAASGG 85  
|||||  
QY 15 AAASGG 20

RESULT 6  
ID P7139 PRELIMINARY; PRT; 115 AA.  
AC P7139;  
DT 01-FEB-1997 (TREMREL. 02, CREATED)  
DT 01-FEB-1997 (TREMREL. 02, LAST SEQUENCE UPDATE)  
DE RTN (FRAGMENT).  
GN RTN  
OS ESCHERICHIA COLI.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RA HWANG U.W., LEE D.W., KIM C.S., CHAE K.S.;  
RL MOL. CELLS 4:387-391(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE; 97234659.  
RA HALL B.G.;  
RT "The rtn gene of Proteus vulgaris is actually from Escherichia coli."  
RL J. BACTERIOL. 179:2433-2434(1997).  
DR EMBL: U83404; G1785617; -  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 13152 MW; 0C2781BD CRC32;

Query Match 30.0%; Score 6; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 58 TLAKRL 63  
|||||  
QY 8 TLAKRL 13

RESULT 7  
ID Q24413 PRELIMINARY; PRT; 120 AA.  
AC Q24413;  
DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE ACIDIC RIBOSOMAL PROTEIN P3A.  
GN RPP3A.  
OS ZEA MAYS (MAIZE).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;  
OC POACEAE; ZEA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B73; TISSUE=IMMATURE EAR;  
RX MEDLINE; 97422884.  
RA BAILEY-SERRES J., VANGALA S., SZICK K., LEE C.H.;  
RT "Acidic phosphoprotein complex of the 60S ribosomal subunit of maize seedling roots. Components and changes in response to flooding."  
RL PLANT PHYSIOL. 114:1293-1305(1997).  
DR EMBL: U62751; G2431767; -  
DR PFAM: PF00428; 60s\_ribosomal; 1.  
KW RIBOSOMAL PROTEIN.  
SQ SEQUENCE 120 AA; 12219 MW; D724A261 CRC32;

Query Match 30.0%; Score 6; DB 10; Length 120;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 87 AAASGG 92  
|||||  
QY 15 AAASGG 20

RESULT 8  
ID P70655 PRELIMINARY; PRT; 132 AA.  
AC P70655;  
DT 01-FEB-1997 (TREMREL. 02, CREATED)  
DT 01-FEB-1997 (TREMREL. 02, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE PBX1-HOMEOBOX PROTO-ONCOGENE (FRAGMENT).  
GN PBX1.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 96194581.  
RA REDMOND L., HOCKFIELD S., MORABITO M.A.;  
RT "The divergent homeobox gene PBX1 is expressed in the postnatal subventricular zone and interneurons of the olfactory bulb."  
RL J. NEUROSCI. 16:2972-2982(1996).  
DR EMBL: S82001; E258430; -  
KW HOMEOBOX; DNA-BINDING; NUCLEAR PROTEIN.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 14299 MW; 83AFDF48 CRC32;

Query Match 30.0%; Score 6; DB 11; Length 132;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 90 AAASGG 95  
|||||

QY 15 AAASGG 20

## RESULT 9

ID Q17038 PRELIMINARY; PRT: 155 AA.  
AC Q17038;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE SERINE PROTEINASE (FRAGMENT).  
OS ANOPHELES GAMBIAE (AFRICAN MALARIA MOSQUITO).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; NEMATODERA; CULICOIDEA; CULICIDAE; ANOPHELES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-G3;  
RA SIDEN-KIAMOS I., SKAVDIS G., RUBIO J., PAPAGIANNAKIS G., LOUIS C.;  
RL INSECT MOL. BIOL. 5:61-71(1996).  
DR EMBL; Z49832; G995967; -.  
DR PFAM; PF00089; trypsin; 1.  
FT NON\_TER 155 155  
SQ SEQUENCE 155 AA; 17118 MW; 3BD46B1B CRC32;

Query Match 30.0%; Score 6; DB 5; Length 155;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 AKRLPA 126

|||||

QY 10 AKRLPA 15

## RESULT 10

ID O46550 PRELIMINARY; PRT: 162 AA.  
AC O46550;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE CAVBOLIN-2.  
GN CAV-2.  
OS CANIS FAMILIARIS (DOG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC CARNIVORA; FISSIPEDIA; CANIDAE; CANIS.  
RN [1]  
RA SCHEFFELE P., VERKADE P., FRA A.M., SWEET-VIRTA H., SIMONS K.,  
RA IKONEN E.;  
RL J. CELL BIOL. 0:0-0(1998).  
DR EMBL; AF039223; G2724139; -.  
SQ SEQUENCE 162 AA; 17871 MW; 1289658C CRC32;

Query Match 30.0%; Score 6; DB 6; Length 162;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 36 SGSDRD 41

|||||

QY 2 SGSDRD 7

## RESULT 11

ID O00439 PRELIMINARY; PRT: 202 AA.  
AC O00439;  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 21.7 KD PROTEIN.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.

RC\*, TISSUE=FRONTAL CORTEX;

RA ALWAZZAN M., HAMSHERE M.G., LENNON G., BROOK J.D.;  
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; Y10936; E305308; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 202 AA; 21669 MW; 553E3EB4 CRC32;

Query Match 30.0%; Score 6; DB 4; Length 202;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 RLPAAA 24

|||||

QY 12 RLPAAA 17

## RESULT 12

ID O91255 PRELIMINARY; PRT: 203 AA.  
AC O91255;  
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE COAT PROTEIN  
OS PEA EARLY BROWNING VIRUS.  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; TOBRNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DUTCH;  
RA SWANSON M.M., MACFARLANE S.A.;  
RT "The Ell6 isolate of Dutch pea early-browning virus is a  
RT recombinant."  
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AJ006500; E1318178; -.  
KW COAT PROTEIN.  
SQ SEQUENCE 203 AA; 21857 MW; 244A85DB CRC32;

Query Match 30.0%; Score 6; DB 14; Length 203;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 127 PAAASG 132

|||||

QY 14 PAAASG 19

## RESULT 13

ID Q12097 PRELIMINARY; PRT: 204 AA.  
AC Q12097;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 22.6 KD PROTEIN (FDS PROTEIN).  
OS FLAMMULINA VELUTIPES.  
OC EUKARYOTA; FUNGI; BASIDIOMYCOTA; HYMENOMYCETES; AGARICALES;  
OC TRICHOLOMATACEAE; FLAMMULINA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE 4;  
RA AZUMA T., HARADA A., KIM D., SAKUMA Y., KOJIMA Y., MIURA K.;  
RL MOKUZAI GAKKAISHI 42:688-692(1996).  
DR EMBL; D83744; D1012768; -.  
DR EMBL; D83658; D1012701; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 204 AA; 22628 MW; 48EB0BDD CRC32;

Query Match 30.0%; Score 6; DB 3; Length 204;  
Best Local Similarity 100.0%; Pred. No. 2.53e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 SGSDRD 156

|||||

QY 2 SGSDRD 7

Search completed: Sat Aug 28 14:48:19 1999  
Job time : 49 secs.

```

RESULT 14
ID Q22687 PRELIMINARY; PRT; 206 AA.
AC Q22687;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)
DE T22H6.1 PROTEIN.
GN T22H6.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA MCMURRAY A.;
RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL NATURE 368:32-38(1994).
DR EMBL: 250797; E1349640; -.
SQ SEQUENCE 206 AA; 23959 MW; E465429E CRC32;

Query Match 30.0%; Score 6; DB 5; Length 206;
Best Local Similarity 100.0%; Pred.No. 2.53e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 128 KRLPAA 133
QY 11 KRLPAA 16

RESULT 15
ID Q01400 PRELIMINARY; PRT; 209 AA.
AC Q01400;
DT 01-JUL-1997 (TREMREL. 04, CREATED)
DT 01-JUL-1997 (TREMREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE BLASTODERM-SPECIFIC PROTEIN 25A.
GN BSG25A.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97473535.
RA SINGER J.B., LENGUEL J.A.;
RT "Expression and sequence analysis of the Drosophila
RT blastoderm-specific gene bsg25A."
RL GENE 197:379-382(1997).
DR EMBL: U80801; G1930012; -.
DR FLYBASE: FBgn0000227; Bsg25A.
SQ SEQUENCE 209 AA; 23582 MW; A317B6FA CRC32;

Query Match 30.0%; Score 6; DB 5; Length 209;
Best Local Similarity 100.0%; Pred.No. 2.53e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 134 DRDTLA 139
QY 5 DRDTLA 10

```

**This Page Blank (uspto)**

\*\*\*\*\*

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Run on: Mon Aug 23 13:09:20 1999; MasPar time 15.52 Seconds  
Tabular output not generated. 27.407 Million cell updates/sec

Title: >US-09-049-696-44  
Description: (1-20) from US09049696.pep  
Perfect Score: 120  
Sequence: 1 NSGSDRTLAKRLPAAASGG 20

Scoring table: PAM 150  
Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 19.420; Variance 64.413; scale 0.301

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	58	48.3	530	34	W49042 Human low density lip	6.74e+01
2	58	48.3	557	34	W49039 Rabbit low density li	6.74e+01
3	56	46.7	349	29	W44368 Aspergillus nidulans	1.05e+02
4	55	46.7	517	1	R06260 Human acetylcholine r	1.05e+02
5	55	45.8	28	6	R32568 CBH I hinge linker pe	1.31e+02
6	55	45.8	28	4	R21408 Sequence of a single	1.31e+02
7	55	45.8	269	6	R32569 Fusion protein encode	1.31e+02
8	55	45.8	354	24	W44369 Aspergillus nidulans	1.31e+02
9	55	45.8	513	14	R77261 T. longibrachiatum ex	1.31e+02
10	55	45.8	513	3	P50133 Sequence encoded by c	1.31e+02
11	55	45.8	513	18	W02022 Trichoderma cellbioh	1.31e+02
12	55	45.8	1307	19	R99255 Aspergillus flavus mu	1.31e+02
13	54	45.0	24	6	R32566 Peptide encoded by CB	1.63e+02
14	54	45.0	197	2	P70243 AA sequence (III) of	1.63e+02
15	54	45.0	197	1	P80478 Sequence of a polypep	1.63e+02
16	54	45.0	197	15	R93136 Human guanylate kinas	1.63e+02

17	54	45.0	440	4	R22362 GroEL-1 protein parti	1.63e+02
18	54	45.0	540	4	GroEL-1 protein.	1.63e+02
19	54	45.0	548	18	R94368 Brevibacterium flavum	1.63e+02
20	54	45.0	686	34	W64469 Human secreted protei	1.63e+02
21	54	45.0	1013	18	R99642 Poly(ADP-ribose) poly	1.63e+02
22	54	45.0	2799	39	W81867 Human tumour suppress	1.63e+02
23	54	44.2	309	31	W60146 M. vaccae antigen GV-	2.03e+02
24	53	44.2	523	31	W60144 M. vaccae antigen GV-	2.03e+02
25	53	44.2	540	2	P81351 Sequence of Mycobacte	2.03e+02
26	53	44.2	540	29	W44702 Mycobacterium tubercu	2.03e+02
27	53	44.2	540	27	W32100 Mycobacteria sp. heat	2.03e+02
28	53	44.2	540	16	R81610 Mycobacterium tubercu	2.03e+02
29	53	44.2	541	12	R67384 M. leprae GroEL gene	2.03e+02
30	53	44.2	544	27	W32099 Mycobacteria sp. heat	2.03e+02
31	53	44.2	560	2	P80215 Sequence of Mycobacte	2.03e+02
32	53	44.2	572	1	R04716 Amino acid sequence o	2.03e+02
33	53	44.2	573	12	R64765 M. leprae 65 kDa prot	2.03e+02
34	53	44.2	573	12	R64766 M. tuberculosis 65 kD	2.03e+02
35	53	44.2	573	1	R04715 Amino acid sequence o	2.03e+02
36	53	44.2	588	1	P80364 M. leprae 65kD antige	2.03e+02
37	53	44.2	1078	13	R71704 Collagen alpha 1 (III	2.03e+02
38	53	44.2	1383	33	W48842 Human receptor tyrosi	2.03e+02
39	52	43.3	520	1	R06262 Calf acetylcholine re	2.51e+02
40	52	43.3	527	2	R05079 Human monoamine oxida	2.51e+02
41	52	43.3	527	33	W61278 Monamine oxidase A.	2.51e+02
42	52	43.3	621	7	R37870 Soluble KEX2 protease	2.51e+02
43	52	43.3	622	4	R20469 Soluble yscF.	2.51e+02
44	52	43.3	712	4	R20468 KEX2 endopeptidase wi	2.51e+02
45	52	43.3	814	27	W25685 Kex2 protease.	2.51e+02

ALIGNMENTS

RESULT 1  
ID W49042 standard; Protein: 530 AA.  
AC W49042;  
DT 09-NOV-1998 (first entry)  
DE Human low density lipoprotein binding protein LBP-3.  
KW Low density lipoprotein binding protein; LDL binding protein 3;  
LBP-3; receptor; human; atherosclerosis; diagnosis; therapy;  
KW vaccine.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 53..59  
FT /note= "Claim 2"

PN WO9823282-A1.  
PD 04-JUN-1998.  
PF 26-NOV-1997; U21857.  
PR 03-JUN-1997; US-048547.  
PR 27-NOV-1996; US-031930.  
PA (BOST-) BOSTON HEART FOUND INC.  
PI Arjona AA, Law SW, Lees AM, Lees RS;  
DR WPI: 98-322455/28.  
DR N-PSDB: V32839.  
PT Nucleic acid encoding low density lipoprotein binding proteins and  
related vectors - transformed cells, proteins, and modulators of  
PT binding, useful for treatment and diagnosis of atherosclerosis and  
PT for identifying subjects at risk  
PS Claim 1; Fig 8; 47pp; English.  
CC This polypeptide comprises novel human low density lipoprotein  
CC (LDL) binding protein LBP-3 that is capable of binding both native  
CC and methyl LDL. Its amino acid sequence was deduced from an  
CC isolated cDNA clone (see V32839). cDNA clones (see V32834-39)  
CC and encoded rabbit and human LBPs (see W49037-42) are claimed. An  
CC abnormality in an aspect of LBP metabolism or structure is  
CC diagnostic of a risk for atherosclerosis. The invention provides:  
CC methods for determining if an animal is at risk for atherosclerosis  
CC (e.g. for prenatal screening); methods for treating atherosclerosis  
CC (including gene therapy) using e.g. LBP polypeptides to bind LDL and  
CC thereby prevent formation of atherosclerotic plaque; and methods  
CC for treating a cell having an abnormality in LBP structure or  
CC metabolism. Pharmaceutical and vaccine compositions are also  
CC provided, as well as recombinant vectors and host cells used to

CC produce recombinant LBP.  
SQ Sequence 530 AA;

Query Match 48.3%; Score 58; DB 34; Length 530;  
Best Local Similarity 35.3%; Pred. No. 6.74e+01;  
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 460 terndlnkrvqdsagg 476  
:::|:|:|:|:|:|

Qy 4 SDRDTLAKRLPAAASGG 20

RESULT 2  
ID W49039 standard; Protein; 557 AA.  
AC W49039;  
DE 09-NOV-1998 (first entry)  
KW Rabbit low density lipoprotein binding protein LBP-3.  
KW Low density lipoprotein binding protein; LDL binding protein 3;  
KW LBP-3; receptor; rabbit; atherosclerosis; diagnosis; therapy;  
KW vaccine.  
OS Oryctolagus cuniculus.  
FH Key Location/Qualifiers  
FT Peptide 96..110  
FT /note="Claim 2"  
PN W09823282-A1.  
PD 04-JUN-1998.  
PF 26-NOV-1997; U21857.  
PR 03-JUN-1997; US-048547.  
PR 27-NOV-1996; US-031930.  
PA (BOST-) BOSTON HEART FOUND INC.  
PI Arjona AA, Law SW, Lees AM, Lees RS;  
WI: 98-322455/28.  
DR N-PSDB; V32836.  
PT Nucleic acid encoding low density lipoprotein binding proteins and related vectors - transformed cells, proteins, and modulators of binding, useful for treatment and diagnosis of atherosclerosis and for identifying subjects at risk  
PS Claim 1; Fig 5; 47pp; English.  
CC This polypeptide comprises novel rabbit low density lipoprotein (LDL) binding protein LBP-3 that is capable of binding both native and methyl LDL. Its amino acid sequence was deduced from rabbit abdominal aorta cDNA (see V32836). cDNA clones (see V32834-39) and undecoded rabbit and human LBPs (see W49037-42) are claimed. An abnormality in an aspect of LBP metabolism or structure is diagnostic of a risk for atherosclerosis. The invention provides methods for determining if an animal is at risk for atherosclerosis (e.g. for prenatal screening); methods for treating atherosclerosis (including gene therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent formation of atherosclerotic plaque; and methods for treating a cell having an abnormality in LBP structure or metabolism. Pharmaceutical and vaccine compositions are also provided, as well as recombinant vectors and host cells used to produce recombinant LBP.  
SQ Sequence 557 AA;

Query Match 48.3%; Score 58; DB 34; Length 557;  
Best Local Similarity 35.3%; Pred. No. 6.74e+01;  
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 476 terndlnkrvqdsagg 492  
:::|:|:|:|:|:|

Qy 4 SDRDTLAKRLPAAASGG 20

RESULT 3  
ID W44368 standard; Protein; 349 AA.  
AC W44368;  
DE 29-MAY-1998 (first entry)  
KW Aspergillus nidulans metallo-protease pepH  
KW Aspergillus nidulans; metallo-protease; pepH; protein degradation; fungus; food processing.  
OS Aspergillus nidulans.  
PN W09746689-A1.

PD 11-DEC-1997. E02982.  
PF 05-JUN-1997; EP-201579.  
PR 05-JUN-1996; EP-201579.  
PA (KONN ) GIST-BROCADES BV.  
PI Van Den Hombergh JPTW, Visser J;  
DR WPI: 98-042197/04.  
DR N-PSDB: V15305, V15306;  
PT Metallo-protease deficient fungus with site selected DNA disruption  
PT - and Aspergillus metallo-protease genes, useful in protein  
PT production to reduce protease activity hence protein degradation  
PS Example 4; Page 26-27; 53pp; English.  
CC The present sequence represents a metallo-protease, pepH, from  
CC Aspergillus nidulans from the present invention. The present invention  
CC describes a new protease deficient filamentous fungus, optionally with  
CC reduced extracellular acid protease activity, containing a site selected  
CC disruption of DNA resulting in reduced metallo-protease activity. The  
CC fungi are useful for the production of (heterologous and homologous)  
CC proteins e.g. for food processing, since reduced protease activity  
CC minimises the chance that, and rate at which, the proteins are degraded  
CC during production. DNA sequences encoding metallo-proteases can be  
CC used to produce metallo-protease deficient fungi, by transforming a  
CC filamentous fungus mutant with the constructs and selecting a  
CC transformed fungus with reduced metallo-protease activity. They are also  
CC useful for producing filamentous fungal metallo-protease, by culturing  
CC filamentous fungi transformed with the constructs under suitable  
CC conditions for sequence expression and recovering the metallo-protease.  
CC Such metallo-proteases are useful to assess in vitro whether proteins  
CC which it is proposed to produce from a fungal host are susceptible to  
CC the protease, so determining which metallo-protease genes need to be  
CC inactivated in the host. They are also useful in industrial processes.  
SQ Sequence 349 AA;

Query Match 46.7%; Score 56; DB 29; Length 349;  
Best Local Similarity 50.0%; Pred. No. 1.05e+02;  
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 219 dsststretvaarlkavaa 236  
:|:|:|:|:|:|:|:|:|

Qy 1 NSGSDRDTLAKRLPAAAS 18

RESULT 4  
ID R06260 standard; protein; 517 AA.  
AC R06260;  
DE 07-DEC-1990 (first entry)  
DE Human acetylcholine receptor TE671 (Achr) delta-subunit.  
KW Nicotinic acetyl choline receptor; Achr; TE671; insecticides;  
KW Muscle relaxants; anthelmintics.  
OS Homo sapiens.  
PN CA2003459-A.  
PD 23-MAY-1990.  
PF 21-NOV-1989; 003459.  
PR 23-NOV-1988; US-275422.  
PA (SALK ) SALK INST FOR BIOL STUD.  
PI Lindstrom JM, Schoepfer RD;  
DR WPI: 90-231525/31.  
DR N-PSDB; Q05558.  
PT Human muscle nicotinic acetyl:choline receptor - used to assay  
PT the effects of agents which affect acetyl:choline receptors in  
PT skeletal muscles.  
PS Disclosure; English.  
CC Receptors may be used in assay for materials which modify them.  
CC They may be produced in substantial, pure quantities for use in  
CC experimentation, development of insecticides without effect on  
CC hMNARS and treatment of parasitic infections. MABs raised to the  
CC peptides may be useful in detection of the structure of MNARS.  
SQ Sequence 517 AA;

Query Match 46.7%; Score 56; DB 1; Length 517;  
Best Local Similarity 53.8%; Pred. No. 1.05e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 405 serhglarlitta 417

```

QY      4 SDRDTLAKRLPAA 16
      1:| | | | | | | |
      2 SDRDTLAKRLPAA 16
      1:| | | | | | | |

RESULT 5
ID R32568 standard; Protein; 28 AA.
AC R32568;
DT 08-JUN-1993 (first entry)
DE CBH I hinge linker peptide.
KW Spacer peptide; secretab; single chain; fusion protein; antibody;
KW scAb; recombinant; rDNA; linker; pML2.
OS Synthetic.
FH Key
FT region 1..22 Location/Qualifiers
FT region /note= "SS"
FT region 23..135
FT region /note= "VH"
FT region 136..163
FT region /note= "CBHI hinge"
FT region 164..277
FT region /note= "VL"
PN F19103434-A. (First Major Country Equivalent).
PN W09302198-A.
PD 17-JAN-1992.
PR 16-JUL-1991; 913434.
PR 16-JUL-1991; US-552751.
PA (TERE-) TECH RES CENT FINLAND.
PI Alfthan K, Knowles JKC, Laukkanen ML, Sizmann D, Takkinen K, Teeri TT;
DR WPI: 92-134225/17.
DR N-PSDB; Q36981.
PT Prod. of single chain fusion protein, pref. antibody - comprises
PT transforming host cells, e.g. E. coli with expression constructs
PT composed of proteins or domains, linked by spacer peptide(s)
PS Example; Fig 2: 56pp; English.
CC The sequence is that of the CBH I hinge linker peptide which may
CC be used as part of the prodn. of secretab, biologically active
CC single chain antibodies (scAbs) and other secretab fusion
CC proteins having at least 2 distinct functional proteins or domains.
SQ Sequence 28 AA;

Query Match 45.8%; Score 55; DB 6; Length 28;
Best Local Similarity 36.8%; Pred. No. 1.31e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 1 ppgnrgtttrtpatttgs 19
      1:| | | | | | | |
QY 2 SDRDTLAKRLPAA 20
      1:| | | | | | | |

RESULT 6
ID R21408 standard; peptide; 28 AA.
AC R21408;
DT 17-MAY-1992 (first entry)
DE Sequence of a single chain antibody, having cbh1 hinge as a linker
DE encoded by pML5.
KW Immunoglobulin; fusion protein; heavy chain; light chain.
PN W09201797-A.
PD 06-FEB-1992.
PR 15-JUL-1991; F10218.
PR 16-JUL-1990; US-552757.
PA (ALKO-) ALKO OY.
PI Nyssönen E, Keranen S, Penttilä M, Takkinen K, Knowles JKC;
DR WPI: 92-064953/08.
DR N-PSDB; Q21395.
PT Immunoglobulin gene-contg. vector for transformation of
PT trichoderma - recombinant proteins useful in diagnosis and
PT control of physiological and agricultural conditions
PS Example; Fig 29; 138pp; English.
CC The inventors claim a fusion protein (FP) comprising immunologically
CC active fragments of light and heavy chains linked to each other by
CC synthetic, trichoderma derived or heterologous linkers. The FP
CC comprises a variable domain of the light chain or heavy chain. The
CC FPs are useful in diagnosis and control of physiological and
CC agricultural conditions.
SQ Sequence 28 AA;

Query Match 45.8%; Score 55; DB 4; Length 28;
Best Local Similarity 36.8%; Pred. No. 1.31e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 1 ppgnrgtttrtpatttgs 19
      1:| | | | | | | |

QY 2 SGRDRTLAKRLPAA 20
      1:| | | | | | | |

RESULT 7
ID R32569 standard; Protein; 269 AA.
AC R32569;
DT 08-JUN-1993 (first entry)
DE Fusion protein encoded by Ox VH-hinge-VL insert.
KW Spacer peptide; secretab; single chain; antibody; recombinant;
KW scAb; rDNA; linker; Bos taurus.
OS Synthetic.
FH Key
FT region 1..22 Location/Qualifiers
FT region /note= "SS"
FT region 23..135
FT region /note= "VH"
FT region 136..163
FT region /note= "CBHI hinge"
FT region 164..277
FT region /note= "VL"
PN F19103434-A. (First Major Country Equivalent).
PN W09302198-A.
PD 17-JAN-1992.
PR 16-JUL-1991; 913434.
PR 16-JUL-1991; US-552751.
PA (TERE-) TECH RES CENT FINLAND.
PI Alfthan K, Knowles JKC, Laukkanen ML, Sizmann D, Takkinen K, Teeri TT;
DR WPI: 92-134225/17.
DR N-PSDB; Q36982.
PT Prod. of single chain fusion protein, pref. antibody - comprises
PT transforming host cells, e.g. E. coli with expression constructs
PT composed of proteins or domains, linked by spacer peptide(s)
PS Example; Fig 4: 56pp; English.
CC The sequence is that of the fusion protein encoded by the Ox
CC VH-CBH1 hinge-VL insert which was used as part of a method for
CC cloning secretab, biologically active single chain antibodies
CC (scAbs) and other secretab fusion proteins having at least 2
CC distinct functional proteins or domains.
SQ Sequence 269 AA;

Query Match 45.8%; Score 55; DB 6; Length 269;
Best Local Similarity 36.8%; Pred. No. 1.31e+02;
Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 136 ppgnrgtttrtpatttgs 154
      1:| | | | | | | |
QY 2 SGRDRTLAKRLPAA 20
      1:| | | | | | | |

RESULT 8
ID W44369 standard; Protein; 354 AA.
AC W44369;
DT 29-MAY-1998 (first entry)
DE Aspergillus nidulans metallo-protease pepI.
KW Aspergillus nidulans; metallo-protease; pepI; protein degradation;
KW fungus; food processing.
OS Aspergillus nidulans.
FH Key
FT region 1..22 Location/Qualifiers
FT region /note= "encoded by GTC"
PN W09746689-A1.
PD 11-DEC-1997.
PR 05-JUN-1997; E02982.
PR 05-JUN-1996; EP-201579.
PA (KONN ) GIST-BROCADES BV.
PI Van Den Hombergh-JPTW, Visser J;
DR WPI: 98-042197/04.
DR N-PSDB; V15307, V15308.
PT Metallo-protease deficient fungus with site selected DNA disruption
PT - and Aspergillus metallo-protease genes, useful in protein
PT production to reduce protease activity hence protein degradation
PS Example 5; Page 29-30; 53pp; English.
SQ The present sequence represents a metallo-protease, pepI, from

```

CC Aspergillus nidulans from the present invention. The present invention  
 CC describes a new protease deficient filamentous fungus, optionally with  
 CC reduced extracellular acid protease activity, containing a site selected  
 CC disruption of DNA resulting in reduced metallo-protease activity. The  
 CC fungi are useful for the production of (heterologous and homologous)  
 CC proteins e.g. for food processing, since reduced protease activity  
 CC minimises the chance that, and rate at which, the proteins are degraded  
 CC during production. DNA sequences encoding metallo-proteases can be  
 CC used to produce metallo-protease deficient fungi, by transforming a  
 CC filamentous fungus mutant with the constructs and selecting a  
 CC transformed fungus with reduced metallo-protease activity. They are also  
 CC useful for producing filamentous fungal metallo-protease, by culturing  
 CC filamentous fungi transformed with the constructs under suitable  
 CC conditions for sequence expression and recovering the metallo-protease.  
 CC Such metallo-proteases are useful to assess in vitro whether proteins  
 CC which it is proposed to produce from a fungal host are susceptible to  
 CC the protease, so determining which metallo-protease genes need to be  
 CC inactivated in the host. They are also useful in industrial processes.  
 CC Sequence 354 AA;

Query Match 45.8%; Score 55; DB 29; Length 354;  
 Best Local Similarity 50.0%; Pred. No. 1.31e+02;  
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 224 dsstnavaarilravas 241  
 QY 1 NSGSDRDTLAKRLPAAAS 18  
 :|:| |::| | | | |

## RESULT 9

ID R77261 standard; Protein; 513 AA.

AC R77261;

DT 13-DEC-1995 (first entry)

DE T. longibrachiatum exo-cellobiohydrolase CBHI.

KW Cellulase; cellulose; signal; catalytic core; cellulase binding;

OS Trichoderma longibrachiatum.

FH Key Location/Qualifiers

FT peptide

FT 1..18

FT /label= signal

FT /note= "SEQ ID 26"

FT domain

FT 19..458

FT /label= catalytic core

FT /note= "SEQ ID 10"

FT region

FT 459..482

FT /label= linker

FT /note= "SEQ ID 18"

FT domain

FT 483..513

FT /label= binding domain

FT /note= "SEQ ID 2"

FT WO9516782-A.

PN 22-JUN-1995.

PD 19-DEC-1994; 014163.

PR 17-DEC-1993; US-169948.

PA (GENV ) GENENCOR INT INC.

PI Clarkson KA, Collier KD, Fowler T, Larenas E, Ward M;

DR WPI: 95-231574/30.

DR N-PSDB: Q91281.

PT Pure, truncated fungal cellulase protein from Trichoderma - useful to

PT reduce or eliminate dye, colourant or pigment back-staining or

PT redeposition in stone-washing or bio-polishing

PS Claim 11; Figure 1; 105pp; English.

CC Figure 1 depicts the genomic DNA and AA sequence of CBHI derived

CC from T. longibrachiatum. A truncated fungal cellulase protein in

CC which the CBHI catalytic core consists essentially of the AA

CC sequence in SEQ ID no 2 is claimed. Genes for CBHI has been isolated

CC from T. longibrachiatum and the protein domain structure has been

CC confirmed (Shoemaker, S. et al. 1983, Bio/Technology, 1, 691-696;

CC Teeri, T. et al. 1983, Bio/Technology 1, 696-699 and Teeri, T. et

CC al., 1987, Gene, 51, 43-52). DNA gene fragments encoding SEQ IDs

CC 10; 10 and 18; and 10, 18 and 14 are also claimed.

CC Sequence 513 AA;

Query Match 45.8%; Score 55; DB 14; Length 513;  
 Best Local Similarity 36.8%; Pred. No. 1.31e+02;  
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 455 pgngnrgtttrrrpattgs 473  
 QY 2 SGSDRDTLAKRLPAAASG 20  
 :|:| | | | | | | | |

## RESULT 10

ID P50133 standard; Protein; 513 AA.

AC P50133;

DT 30-NOV-1991 (first entry)

DE Sequence encoded by cellobiohydrolase-I (CBHI) gene.

KW Degradation; enzyme; cellulose; recycle; hydrolysis.

OS Trichoderma reesei strain L27.

FH Key Location/Qualifiers

FT peptide

FT 1..17

FT /label= signal

FT 18..513

PN EP-137280-A.

PD 17-APR-1985.

PF 29-AUG-1984; 110305.

PR 31-AUG-1983; US-528216.

PR 16-JUL-1984; US-630374.

PA (CETU ) CETUS CORP.

PI Shoemaker SP, Gelfand DH, Innis MA, Kwok SY, Ladner MB,

PI Schweickart V;

DR N-PSDB: NS0178.

PT Glycosylated fungal cellulase(s) prodn. in large amounts - by

PT recombinant DNA techniques with fungal-derived cellulase gene

PS Disclosure; Table I, Page 39-44; 84pp; English.

CC The inventors claim CBHI and EGI from T.reesei, and recombinant DNA

CC sequences encoding them, free of introns. The enzymes are useful in

CC the degradation of cellulose. They may also be used for controlled

CC and limited hydrolysis of food materials, e.g. for improving the

CC flavour of coffee.

CC Sequence 513 AA;

Query Match 45.8%; Score 55; DB 3; Length 513;  
 Best Local Similarity 36.8%; Pred. No. 1.31e+02;  
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 455 pgngnrgtttrrrpattgs 473  
 QY 2 SGSDRDTLAKRLPAAASG 20  
 :|:| | | | | | | | |

## RESULT 11

ID W02022 standard; Protein; 513 AA.

AC W02022;

DT 28-OCT-1996 (first entry)

DE Trichoderma cellobiohydrolase I.

KW Cellobiohydrolase I; CBHI; cellulase; cellulose; denim;

KW stonewashing; dye redeposition; backstaining.

OS Trichoderma longibrachiatum.

FH Key Location/Qualifiers

FT peptide

FT 1..17

FT /label= Sig\_peptide

FT 18..513

FT /label= Mat\_protein

FT domain

FT 18..458

FT /label= Catalytic\_core\_domain

FT /note= "catalytic core domain is the preferred

FT domain for use in constructs of the

FT invention"

FT region

FT 459..482

FT /label= Linker\_region

FT domain

FT 483..513

FT /label= Cellulose\_binding\_domain

PN WO9623928-A1.



PD 08-AUG-1996.  
 PF 29-JAN-1996; U00977.  
 PR 01-FEB-1995; US-382452.  
 PA (GEMV) GENENCOR INT INC.  
 PI Clarkson KA, Collier KD, Fowler T, Larenas E, Ward M;  
 DR WPI; 96-371466/37.  
 DR N-PSDB; T32220.  
 PT Treatment of cellulose-contg. fabrics such as denim, e.g.  
 PT stone-washing - using truncated cellulase enzyme to increase  
 PT abrasion and give reduced redeposition of dye  
 PS Disclosure; Fig 1A-1E; 124pp; English.  
 CC The amino acid sequences for *Trichoderma longibrachiatum*  
 CC cellobiohydrolase I (CBHI) (W02022), CBHI (W02025), endoglucanase I  
 CC (EGI) (W02029), EGI (W02032) and EGI (W02034) were deduced from  
 CC the respective genomic DNA sequences (T32220-24). The CBHI, CBHI,  
 CC EGI and EGI enzymes have catalytic core domains useful for reducing  
 CC dye redeposition (backstaining) on cellulose-contg. fabrics such as  
 CC denim, whilst maintaining or increasing abrasion during stonewashing.  
 CC Truncated enzymes comprising these catalytic core domains can be obtd.  
 CC by proteolysis of the complete enzyme or by inserting the appropriate  
 CC DNA fragment into a vector, using this to transform a *Trichoderma*  
 CC sp. host cell, and recovering the recombinant core domain.  
 SQ Sequence 513 AA;

Query Match 45.8%; Score 55; DB 18; Length 513;  
 Best Local Similarity 36.8%; Pred. No. 1.31e+02;  
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 455 ppgnrgtttrrrpatttgs 473  
 :||| : |||:  
 QY 2 SGSDRDTLAKRLPAAASGG 20

## RESULT 12

ID R99255 standard; Protein; 1307 AA.  
 AC R99255;  
 DT 01-DEC-1996 (first entry)  
 DE Aspergillus flavus multidrug resistance protein Afl-MDR1.  
 KW Multiple drug-resistance; MDR; Afl-MDR1; antifungal.  
 OS Aspergillus flavus.  
 PN WO9626952-A1.  
 PD 06-SEP-1996.  
 PF 23-FEB-1996; U02665.  
 PR 27-FEB-1995; US-395246.  
 PA (ELIL) LILLY & CO ELI.  
 PI Peery RB, Skatrud PL;  
 DR WPI; 96-412733/41.  
 DR N-PSDB; T35230.  
 PT New multiple drug resistance gene of *Aspergillus flavus* - used to  
 PT identify cpds. that inhibit fungal multiple drug resistance activity  
 PS Disclosure; Page 22-27; 37pp; English.  
 CC A new multiple drug resistance (MDR) protein (R99255), designated  
 CC Afl-MDR1, of *Aspergillus flavus* was identified as the product of  
 CC a cDNA clone isolated from a mycelial cDNA library of *A. flavus*.  
 CC Afl-MDR1 can be synthesised by host cells, partic. *Saccharomyces*  
 CC cerevisiae, transformed with vectors that provide for the expression  
 CC of DNA encoding Afl-MDR1. Transformed host cells can be used to  
 CC identify cpds. that inhibit fungal MDR activity.  
 SQ Sequence 1307 AA;

Query Match 45.8%; Score 55; DB 19; Length 1307;  
 Best Local Similarity 47.1%; Pred. No. 1.31e+02;  
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 3 sdxdlmkpikpsptg 19  
 ||| |||:  
 QY 4 SDRDTLAKRLPAAASGG 20

## RESULT 13

ID R32566 standard; Peptide; 24 AA.  
 AC R32566;  
 DT 08-JUN-1993 (first entry)

DE Peptide encoded by CBH I hinge linker DNA oligomer.  
 KW Spacer peptide; secretable; single chain; fusion protein; antibody;  
 KW scAb; recombinant; rDNA; linker; pML2.  
 OS Synthetic.  
 PN F19103434-A.  
 PN WO9302198-A.  
 PD 17-JAN-1992.  
 PF 16-JUL-1991; 913434.  
 PR 16-JUL-1990; US-552751.  
 PA (FERE-) TECH RES CENT FINLAND.  
 PI Alfthan K, Knowles JKC, Laukkanen ML, Sizmann D, Takkinen K, Teeri TT;  
 DR WPI; 92-134225/17.  
 DR N-PSDB; Q36979.  
 PT Prod. of single chain fusion protein, pref. antibody - comprises  
 PT transforming host cells, e.g. *E. coli* with expression constructs  
 PT composed of proteins or domains, linked by spacer peptide(s)  
 PS Example; Fig 1; 5pp; English.  
 CC The sequence is that encoded by a DNA oligomer used in the construction  
 CC of the vector pML2 which contains a DNA sequence coding for the CBH I  
 CC hinge linker.  
 SQ Sequence 24 AA;

Query Match 45.0%; Score 54; DB 6; Length 24;  
 Best Local Similarity 36.8%; Pred. No. 1.63e+02;  
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 1 ppgnrgtttrrrpatttgs 19  
 :||| : |||:  
 QY 2 SGSDRDTLAKRLPAAASGG 20

## RESULT 14

ID P70243 standard; protein; 197 AA.  
 AC P70243;  
 DT 21-FEB-1991 (first entry)  
 DE AA sequence (III) of a polypeptide having human haematopoietic cell  
 DE growth potentiating factor (HCGPF) activity.  
 KW Autoimmune disease therapy; immunodeficient disease;  
 KW bone marrow transplant.  
 OS Homo sapiens.  
 PN EP-232707-A.  
 PD 19-AUG-1987.  
 PF 07-JAN-1987; 100107.  
 PR 09-JAN-1986; JP-002633.  
 PR 18-DEC-1986; JP-302698.  
 PR 08-JAN-1987; JP-002521.  
 PA (AJIN) AJINOMOTO KK.  
 PI Tadatsugu T, Gen Y, Junji H, Shinsuke T, Hiroshi M,  
 PI Nobukazu K;  
 DR WPI; 87-229568/33.  
 DR N-PSDB; N70355.  
 PT Human haematopoietic cell growth potentiating factor - prepd.  
 PT from gene obtd. using RNA from human peripheral blood derived  
 PT mononuclear cells  
 PS Claim 13; Page 62; 101pp; English.  
 CC The HCGPF exhibits immune control and haematopoietic control functions  
 CC over a wide range and may be used in the fields of immunodeficient  
 CC diseases, autoimmune diseases, infectious diseases, hepatitis,  
 CC nephritis, cancers and bone marrow transplantation.  
 SQ Sequence 197 AA;

Query Match 45.0%; Score 54; DB 2; Length 197;  
 Best Local Similarity 50.0%; Pred. No. 1.63e+02;  
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Db 13 sqagkstkllkrllqebhg 30  
 ||| |||:  
 QY 2 SGSDRDTLAKRLPAAASGG 19

## RESULT 15

ID P80478 standard; protein; 197 AA.  
 AC P80478;

DT 15-OCT-1990 (first entry)  
DE Sequence of a polypeptide having human haematopoietic cell growth  
potentiating factor (HCGPF) activity  
KW Human haematopoietic cell growth potentiating factor; immune control;  
haematopoietic control functions; immunodeficient diseases;  
KW autoimmune diseases; infectious diseases.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT region 19..197  
FT /note="A polypeptide with this sequence and a gene  
encoding it are claimed"  
FT region 31..197  
FT /note="As above"  
FN EP-274560-A.  
PD 20-JUL-1988.  
PF 19-JUN-1987; 108782.  
PR 08-JAN-1987; JP-002521,  
PA (AJIN) Ajinomoto Kk.  
PI Taniguchi T, Yamada G, Hamuro J, Taki S, Matsui H, Kashima N;  
DR WPI; 88-199153/29.  
DR N-PSDB; n81032.  
PT Recombinant human haematopoietic cell growth potentiating factor -  
used for immuno-deficient, auto-immune or infectious diseases,  
hepatitis, nephritis, cancer or bone marrow transplantation  
PS Claim 13, claim 34; Fig 2 and Fig 5 and Fig 8; 58pp; English.  
CC A polypeptide with this sequence and a gene encoding it are claimed. Also  
claimed is a vector DNA capable of replication in prokaryotes and  
eukaryotes, the transformed prokaryotes and eukaryotes and a polypeptide  
having HCGPF activity. The HCGPF potentiates the growth of haematopoietic  
cells. It exhibits immune control and haematopoietic control functions.  
SQ Sequence 197 AA;

Query Match 45.0%; Score 54; DB 1; Length 197;  
Best Local Similarity 50.0%; Pred. NO. 1.63e-02;  
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
Db 13 sgagkstillkrllqehsg 30  
||: : || ||| ||  
QY 2 SGSDRDTLAKRLPAAASG 19

Search completed: Mon Aug 23 13:09:45 1999  
Job time : 25 secs.

\*\*\*\*\*

```

#title      Isolation of cDNAs encoding the catalytic domain of poly
            (ADP-ribose)polymerase from Xenopus laevis and cherry
            salmon using heterologous oligonucleotide consensus
            sequences.
#cross-references MUID:93277538
#accession   PNO495
#molecule_type mRNA
##residues   742-745, 'E', 747-876 ##label OZA
COMMENT      This zinc-finger protein plays a role in DNA repair, cell growth,
            and differentiation.
CLASSIFICATION #superfamily NAD+ ADP-ribosyltransferase
KEYWORDS      DNA binding; glycosyltransferase; hexosyltransferase; NAD;
            nucleus; pentosyltransferase; zinc finger
SUMMARY      #length 998 #checksum 6153

Query Match      51.7%; Score 62; DB 2; Length 998;
Best Local Similarity 47.4%; Pred. No. 1.71e+00;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 174 SAEDKSLKKLPAVKNEG 192
QY 2 SGSDRDTLAKRLPAAASGG 20

RESULT 3
ENTRY   S13129 #type complete
TITLE   probable homeotic protein ceh-9 - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE     21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
            17-Oct-1997
ACCESSIONS S13129
REFERENCE   Hawkins, N.C.; McGhee, J.D.
            Nucleic Acids Res. (1990) 18:6101-6106
#journal   Homeobox containing genes in the nematode Caenorhabditis
            elegans.
#title
#cross-references MUID:91045075
#accession S13129
##status   preliminary
##molecule_type DNA
##residues 1-60 ##label HAW
##cross-references EMBL:X52811
GENETICS
#gene      ceh-9
#introns   44/3
CLASSIFICATION #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS      DNA binding; homeobox; nucleus; transcription regulation
FEATURE       #domain homeobox homology #label HOX
SUMMARY      #length 60 #molecular-weight 7308 #checksum 9258

Query Match      50.0%; Score 60; DB 2; Length 60;
Best Local Similarity 52.9%; Pred. No. 3.90e+00;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 27 SSSDRSELAKRLDVTET 43
QY 2 SGSDRDTLAKRLPAAAS 18

RESULT 4
ENTRY   G70602 #type complete
TITLE   hypothetical protein Rv1004c - Mycobacterium tuberculosis
            (strain H37Rv)
ORGANISM #formal_name Mycobacterium tuberculosis
DATE     17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
            17-Jul-1998
ACCESSIONS G70602
REFERENCE   A70500
#authors    Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
            C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry
            III, C.E.; Tekai, F.; Badcock, K.; Basham, D.; Brown, D.;
            Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;

```

```

Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Hornsby, T.; Jagels, K.; Krogh, A.; Mclean, J.; Moule, S.;
Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrell, B.G.
Nature (1998) 393:537-544
Deciphering the biology of Mycobacterium tuberculosis from
the complete genome sequence.
#cross-references MUID:98295987
#accession   G70602
##status     preliminary; nucleic acid sequence not shown;
            translation not shown
##molecule_type DNA
##residues   1-419 ##label COL
##cross-references GB:294752; GB:AL123456; NID:g3261731; PID:e1299781;
            PID:g3261734
##experimental_source strain H37Rv
GENETICS
#gene      Rv1004c
SUMMARY    #length 419 #molecular-weight 38785 #checksum 4947

Query Match      50.0%; Score 60; DB 2; Length 419;
Best Local Similarity 57.9%; Pred. No. 3.90e+00;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 326 AGVDPATLAAALPALAAGG 344
QY 2 SGSDRDTLAKRLPAAASGG 20

RESULT 5
ENTRY   S36742 #type complete
TITLE   cation-transporting ATPase (EC 3.6.1.-) pacL - Synechococcus
            sp.
ORGANISM #formal_name Synechococcus sp.
DATE     10-Dec-1993 #sequence_revision 27-Feb-1997 #text_change
            21-Aug-1998
ACCESSIONS S36742
REFERENCE   S36741
#authors    Kanamaru, K.; Kashiwagi, S.; Mizuno, T.
#journal    FEBS Lett. (1993) 330:99-104
#title      The cyanobacterium, Synechococcus sp. PCC7942, possesses two
            distinct genes encoding cation-transporting P-type ATPases.
#cross-references MUID:93380581
#accession S36742
##status     preliminary
##molecule_type DNA
##residues   1-926 ##label KAN
CLASSIFICATION #superfamily Na+/K+-transporting ATPase alpha chain; ATPase
            nucleotide-binding domain homology
KEYWORDS      hydrolase
FEATURE       #domain ATPase nucleotide-binding domain homology #label
            ATN
SUMMARY      #length 926 #molecular-weight 99696 #checksum 500

Query Match      50.0%; Score 60; DB 2; Length 926;
Best Local Similarity 43.8%; Pred. No. 3.90e+00;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 324 QRESLIRLPVETLIG 339
QY 5 DRDTLAKRLPAAASGG 20

RESULT 6
ENTRY   JH0581 #type complete
TITLE   NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken
            poly(ADP-ribose) synthase
ALTERNATE_NAMES #formal_name Gallus gallus #common_name chicken
ORGANISM #formal_name Gallus gallus
DATE     17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
            08-Sep-1997

```

```

ACCESSIONS      JH0581
REFERENCE        JH0581
#authors         Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.
#journal         Gene (1991) 102:157-164
#title           Chicken poly(ADP-ribose) synthetase: complete deduced amino
                  acid sequence and comparison with mammalian enzyme
                  sequences.
#cross-references EMBL:91340148
#accession       JH0581
#molecule_type mRNA
##residues       1-1011 ##label ITT
##cross-references EMBL:X52690; NID:g63742; PID:g63743
COMMENT          This protein is a chromatin-bound enzyme.
COMMENT          This enzyme catalyzes DNA-dependent post-translational
                  modifications of various nuclear proteins.
CLASSIFICATION  #superfamily NAD+ ADP-riboseyltransferase
KEYWORDS         DNA binding; glycosyltransferase; NAD; nucleus;
                  pentosyltransferase; zinc finger
SUMMARY          #length 1011 #molecular-weight 113605 #checksum 5006

Query Match      50.0%; Score 60; DB 2; Length 1011;
Best Local Similarity 50.0%; Pred.No. 3.90e+00;
Matches          8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 191 DKETLKKQLPATKTEG 206
      ::::|::|::|
Qy 5 DRDTLAKRLPAASGG 20

RESULT 7
ENTRY   C40722 #type fragment
TITLE   homeotic protein ceh-23 - Caenorhabditis elegans (fragment)
ORGANISM #formal_name Caenorhabditis elegans
DATE     12-May-1994 #sequence_revision 19-May-1995 #text_change
          17-Oct-1997

ACCESSIONS      C40722
REFERENCE        A40722
#authors         Wang, B.B.; Mueller-Immergluck, M.M.; Austin, J.; Robinson,
                  N.T.; Chisholm, A.; Kenyon, C.
#journal         Cell (1993) 74:29-42
#title           A homeotic gene cluster patterns the anteroposterior body
                  axis of Caenorhabditis elegans.
#accession       C40722
#status          preliminary; not compared with conceptual translation
#molecule_type mRNA
##residues       1-61 ##label WAN
CLASSIFICATION  #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS         DNA binding; homeobox; nucleus; transcription regulation
FEATURE         2-58
SUMMARY          #domain homeobox homology #label HOX
                  #length 61 #checksum 9879

Query Match      49.2%; Score 59; DB 2; Length 61;
Best Local Similarity 43.8%; Pred.No. 5.84e+00;
Matches          7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 28 GAERENLAORLGLSPS 43
      ::::|::|::|
Qy 3 GSDRDTLAKRLPAAS 18

RESULT 8
ENTRY   S44907 #type complete
TITLE   ZK652.5 protein - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE     14-Sep-1994 #sequence_revision 12-May-1995 #text_change
          17-Oct-1997
ACCESSIONS      S44907
REFERENCE        S44613
#authors         Du, Z.
#submission      submitted to the EMBL Data Library, May 1993
#description      Sequence of the C. elegans cosmid ZK652.
#accession       S44907

```

```

#status          preliminary
#molecule_type DNA
##residues       1-305 ##label DUZ
##cross-references EMBL:L14429; NID:g289765; PID:g289773
GENETICS
#introns         25/1; 82/3; 215/2
CLASSIFICATION  #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS         DNA binding; homeobox; nucleus; transcription regulation
FEATURE         212-268
SUMMARY          #domain homeobox homology #label HOX
                  #length 305 #molecular-weight 33260 #checksum 2605

Query Match      49.2%; Score 59; DB 2; Length 305;
Best Local Similarity 43.8%; Pred.No. 5.84e+00;
Matches          7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 238 GAERENLAORLGLSPS 253
      ::::|::|::|
Qy 3 GSDRDTLAKRLPAAS 18

RESULT 9
ENTRY   C41325 #type complete
TITLE   heat shock protein 56 - Streptomyces albus
ALTERNATE_NAMES heat shock protein groEL homolog 2
ORGANISM #formal_name Streptomyces albus
DATE     17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
          12-Sep-1997

ACCESSIONS      C41325
REFERENCE        A41325
#authors         Mazodier, P.; Guglielmi, G.; Davies, J.; Thompson, C.J.
#journal         J. Bacteriol. (1991) 173:7382-7386
#title           Characterization of the groEL-like genes in Streptomyces
                  albus.
#cross-references EMBL:92041639
#accession       C41325
#molecule_type DNA
##residues       1-540 ##label MAZ
##cross-references GB:M76658; NID:g153293; PID:g153294
GENETICS
#gene            groEL2
CLASSIFICATION  #superfamily chaperonin groEL
SUMMARY          #length 540 #molecular-weight 56717 #checksum 1413

Query Match      49.2%; Score 59; DB 2; Length 540;
Best Local Similarity 42.1%; Pred.No. 5.84e+00;
Matches          8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 355 DSDYDREKLQERLANVAGG 373
      ::|::|::|
Qy 1 NSGSDRDTLAKRLPAASG 19

RESULT 10
ENTRY   A35360 #type complete
TITLE   otefin - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE     17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change
          24-Sep-1998
ACCESSIONS      A35360
REFERENCE        A35360
#authors         Padan, R.; Mainudel-Epszteyn, S.; Goitein, R.; Fainsod, A.;
                  Gruenbaum, Y.
#journal         J. Biol. Chem. (1990) 265:7808-7813
#title           Isolation and characterization of the Drosophila nuclear
                  envelope otefin cDNA.
#cross-references EMBL:90243644
#accession       A35360
#status          preliminary
#molecule_type DNA; mRNA
##residues       1-406 ##label PAD
##cross-references GB:X17495; NID:g8313; PID:g8314
GENETICS

```

```

#gene FlyBase:Ote
#cross-references FlyBase:FBgn0003022
KEYWORDS transmembrane protein
SUMMARY #length 406 #molecular-weight 44658 #checksum 3473

Query Match 48.3%; Score 58; DB 2; Length 406;
Best Local Similarity 44.4%; Pred. No. 8.72e+00;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Db 12 TDSRRKVLVRLRSTGG 29
:|:|:|:|:|:|
Qy 2 SGSDRDTLAKRLPAAASG 19

RESULT 11
ENTRY #type complete
TITLE phenylalanyl-tRNA synthetase, subunit beta (phet) homolog -
ORGANISM Archaeoglobus fulgidus
#formal_name Archaeoglobus fulgidus
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
21-Aug-1998
ACCESSIONS G69427
REFERENCE A69250
#authors Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
K.E.; Ketchum, K.A.; Dodson, R.J.; Swinn, M.; Hickey, E.K.;
Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham,
D.E.; Kyripides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.;
Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
T.; Artach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.;
D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.;
Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
C.R.; Venter, J.C.
#journal Nature (1997) 390:364-370
#title The complete genome sequence of the hyperthermophilic,
sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references GB:AE001005; GB:AE000782; NID:g2689328; PID:g2649146;
TIGR:AF1424
#accession G69427
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-547 #label KLE
#cross-references GB:AE001005; GB:AE000782; NID:g2689328; PID:g2649146;
TIGR:AF1424
CLASSIFICATION #superfamily yeast cytosolic phenylalanine--tRNA ligase alpha
Chain
SUMMARY #length 547 #molecular-weight 62278 #checksum 9769

Query Match 48.3%; Score 58; DB 2; Length 547;
Best Local Similarity 66.7%; Pred. No. 8.72e+00;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 16 GADRTILSRLP 27
:|:|:|:|:|:|
Qy 3 GSDRDTLAKRLP 14

RESULT 12
ENTRY #type complete
TITLE ribosomal protein L27 (rpmA) - Lyme disease spirochete
ORGANISM Borrelia burgdorferi
#formal_name Borrelia burgdorferi #common_name Lyme disease
spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
05-Jun-1998
ACCESSIONS C70197
REFERENCE A70100
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Dodson, R.; Hickey, E.K.; Swinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Utterback, T.; Wathey, L.; McDonald, L.; Artach, P.;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi.
#cross-references MUID:98065943
#accession C70197
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-81 #label KLE
#cross-references GB:AE001177; GB:AE000783; NID:g2688711; PID:g2688720;
TIGR:BB0780
#experimental_source strain B31
CLASSIFICATION #superfamily Escherichia coli ribosomal protein L27;
eubacterial ribosomal protein L27 homology
SUMMARY #length 81 #molecular-weight 8748 #checksum 31

Query Match 47.5%; Score 57; DB 2; Length 81;
Best Local Similarity 45.0%; Pred. No. 1.29e+01;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 9 SSKNGRDSISRLGVKRSRG 28
:|:|:|:|:|:|
Qy 1 NSGSDRDTLAKRLPAAASG 20

RESULT 13
ENTRY #type complete
TITLE heterocyst maturation protein homolog MG065 - Mycoplasma
genitalium (SGC3)
ORGANISM Mycoplasma genitalium
DATE 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change
24-Jul-1998
ACCESSIONS B64207
REFERENCE A64200
#authors Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton,
R.A.; Fleischmann, R.D.; Bult, C.J.; Kerlavage, A.R.;
Sutton, G.; Kelley, J.M.; Fritchman, J.L.; Weidman, J.F.;
Small, K.V.; Sandusky, M.; Fuhrmann, J.; Nguyen, D.;
Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,
J.M.; Tomb, J.F.; Dougherty, B.A.; Bott, K.F.; Hu, P.C.;
Lucier, T.S.; Peterson, S.N.; Smith, H.O.; Hutchison III,
C.A.; Venter, J.C.
#journal Science (1995) 270:397-403
#title The minimal gene complement of Mycoplasma genitalium.
#cross-references MUID:96026346
#accession B64207
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-466 #label TIGR
#cross-references GB:U39686; GB:L43967; NID:g1045737; PID:g1045740;
TIGR:MG065
#experimental_source strain G-37
GENETICS
#genetic_code SGC3
CLASSIFICATION #superfamily unassigned ATP-binding cassette proteins;
ATP-binding cassette homology
ATP: P-loop
KEYWORDS 252-446
FEATURE #domain ATP-binding cassette homology #label ABC\
269-276 #region nucleotide-binding motif A (P-loop)
SUMMARY #length 466 #molecular-weight 54009 #checksum 6438

Query Match 47.5%; Score 57; DB 2; Length 466;
Best Local Similarity 57.1%; Pred. No. 1.29e+01;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 361 DSIQKLPNELSGG 374

```

1:1 1:1 1:1 1:1  
7 DTLAKRLPAAASGG 20

RESULT 14  
ENTRY  
TITLE NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 08-Sep-1997

ACCESSIONS S04200  
REFERENCE S04200  
#authors Huppi, K.; Bhatia, K.; Siwarski, D.; Klinman, D.; Cherney, B.; Smulson, M.  
#journal Nucleic Acids Res. (1989) 17:3387-3401  
#title Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.  
#cross-references MUID:89263780  
#accession S04200  
#molecule\_type DNA  
#residues 1-1013 #label HUP  
#cross-references EMBL:X14206; NID:g49893; PID:g49894

GENETICS  
#map\_position 1  
CLASSIFICATION #superfamily NAD+ ADP-ribosyltransferase  
KEYWORDS DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc finger  
SUMMARY #length 1013 #molecular\_weight 113099 #checksum 5611

Query Match 47.5%; Score 57; DB 2; Length 1013;  
Best Local Similarity 42.1%; Pred. No. 1.29e+01;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 188 SAEDKALKKOLPAKNEG 206  
1:1:1:1:1:1  
QY 2 SGSDRTLAKRLPAAASGG 20

Query Match 46.7%; Score 56; DB 2; Length 353;  
Best Local Similarity 50.0%; Pred. No. 1.91e+01;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 321 RDTIEKRPPSGSAG 334  
1:1:1:1:1:1:1  
QY 6 RDTLAKRLPAAASG 19

Search completed: Mon Aug 23 13:09:00 1999  
Job time : 14 secs.

RESULT 15  
ENTRY  
TITLE C70985 #type complete  
probable pksII protein - Mycobacterium tuberculosis (strain H37RV)  
ORGANISM #formal\_name Mycobacterium tuberculosis  
DATE 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 17-Jul-1998

ACCESSIONS C70985  
REFERENCE A70500  
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

#journal Nature (1998) 393:537-544  
#title Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.  
#cross-references MUID:98295987  
#accession C70985  
#status preliminary; nucleic acid sequence not shown; translation not shown  
#molecule\_type DNA  
#residues 1-353 #label COL  
#cross-references GB:295617; GB:AL123456; NID:g3242249; PID:e1300713; PID:g3242250

#experimental\_source strain H37RV

GENETICS  
#gene pksII  
SUMMARY #length 353 #molecular\_weight 37639 #checksum 6233

**This Page Blank (uspto)**



\*\*\*\*\*

WQSEFH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:07:32 1999; MasPar time 4.36 Seconds  
Tabular output not generated. 129.618 Million cell updates/sec

Title: >US-09-049-696-44  
Description: (1-20) from US09049696.ppe  
Perfect Score: 120  
Sequence: 1 NSGSDRDTLAKRLPAAASGG 20

Scoring table: PAM 150  
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 28.274; Variance 36.012; scale 0.785

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	ID	Description	Pred. No.
1	62	51.7	998	POLY [ADP-RIBOSE] POLY	6.09e-01
2	61	50.8	141	HEMOGLOBIN THEFA-1 CHA	9.57e-01
3	61	50.8	1013	POLY [ADP-RIBOSE] POLY	9.57e-01
4	60	50.0	60	HOMEBOX PROTEIN CEH-9	1.50e+00
5	60	50.0	926	CATION-TRANSPORTING AT	1.50e+00
6	60	50.0	1011	POLY [ADP-RIBOSE] POLY	1.50e+00
7	59	49.2	305	HOMEBOX PROTEIN CEH-2	2.33e+00
8	59	49.2	539	60 KD CHAPERONIN 2 (PR	2.33e+00
9	58	48.3	406	OTEFIN.	3.60e+00
10	57	47.5	81	50S RIBOSOMAL PROTEIN	5.53e+00
11	57	47.5	466	HYPOTHETICAL ABC TRANS	5.53e+00
12	57	47.5	1012	POLY [ADP-RIBOSE] POLY	5.53e+00
13	57	47.5	1302	CELLULOSE SYNTHASE OPE	5.53e+00
14	56	46.7	449	ASPARTATE AMINOTRANSF	8.45e+00
15	56	46.7	516	ACHD_BOVIN	8.45e+00
16	56	46.7	517	ACHD_HUMAN	8.45e+00
17	56	46.7	517	ACHD_RAT	8.45e+00
18	56	46.7	520	ACHD_MOUSE	8.45e+00
19	56	46.7	539	60 KD CHAPERONIN 1 (PR	8.45e+00
20	56	46.7	1201	TRANSCRIPTION-REPAIR C	8.45e+00
21	56	46.7	1442	HYPOTHETICAL 161.2 KD	8.45e+00
22	55	45.8	319	N5,N10-METHENYLTETRAHY	1.28e+01
23	55	45.8	513	EXOGLUCANASE I PRECURS	1.28e+01

24	55	45.8	659	1	HMN1_DROME	HOMEBOX PROTEIN NK-1	1.28e+01
25	55	45.8	1320	1	PUTA_SALTY	PROLINE DEHYDROGENASE	1.28e+01
26	54	45.0	46	1	AOFA_SHEEP	AMINE OXIDASE [FLAVIN-	1.94e+01
27	54	45.0	69	1	AFG3_HUMAN	AFG3-LIKE PROTEIN (EC	1.94e+01
28	54	45.0	196	1	KGUA_HUMAN	GUANYLATE KINASE (EC 2	1.94e+01
29	54	45.0	221	1	CFI_PHAVU	CHALCONE--FLAVONE IS	1.94e+01
30	54	45.0	268	1	PLSX_BACSU	FATTY ACID/PHOSPHOLIP	1.94e+01
31	54	45.0	287	1	VPH_STRVI	VIOMYCIN PHOSPHOTRANSF	1.94e+01
32	54	45.0	527	1	AOFA_BOVIN	AMINE OXIDASE [FLAVIN-	1.94e+01
33	54	45.0	539	1	CHG1_STRCO	60 KD CHAPERONIN 1 (PR	1.94e+01
34	54	45.0	539	1	CHG1_STRAL	60 KD CHAPERONIN 1 (PR	1.94e+01
35	54	45.0	1012	1	POLG_HDVO	STRUCTURAL POLYPEPTIDE	1.94e+01
36	54	45.0	1013	1	POL_HUMAN	POLY [ADP-RIBOSE] POLY	1.94e+01
37	54	45.0	1036	1	AXO1_CHICK	AXONIN-1 PRECURSOR	1.94e+01
38	54	45.0	1530	1	RPCL_TRYBB	DNA-DIRECTED RNA POLYM	1.94e+01
39	54	45.0	1717	1	RPAL_MOUSE	DNA-DIRECTED RNA POLYM	1.94e+01
40	54	45.0	4660	1	LRP2_RAT	LOW-DENSITY LIPOPROTEI	2.90e+01
41	53	44.2	265	1	YQ01_MYCTU	HYPOTHETICAL 27.9 KD P	2.90e+01
42	53	44.2	299	1	Y009_MYCTU	HYPOTHETICAL 31.0 KD P	2.90e+01
43	53	44.2	450	1	A2AA_CAVPO	ALPHA-2A ADRENERGIC RE	2.90e+01
44	53	44.2	540	1	CHG2_MYCLE	60 KD CHAPERONIN 2 (PR	2.90e+01
45	53	44.2	586	1	PYRG_MYCTU	CTP SYNTHETASE (EC 6.3	2.90e+01

ALIGNMENTS

RESULT 1  
ID PPOL\_XENLA STANDARD: PRT: 998 AA.  
AC P31669;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE POLY [ADP-RIBOSE] POLYMERASE (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
DE RIBOSYLTRANSFERASE) (POLY[ADP-RIBOSE] SYNTHETASE) (FRAGMENT).  
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-OVARY;  
RA SAULIER-LE DREAN B.M.;  
RL THESIS (1992), UNIVERSITE DE RENNES, FRANCE.  
RP SEQUENCE OF 742-876 FROM N.A.  
RX MEDLINE; 93277538.  
RA OZAWA Y., UCHIDA K., UCHIDA M., AMI Y., KUSHIDA S., OKADA N.,  
RA MIWA M.;  
RT "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)  
RT polymerase from Xenopus laevis and cherry salmon using heterologous  
RL oligonucleotide consensus sequences";  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 193:119-125(1993).  
CC -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-RIBOSYL)[N]-ACCEPTOR =  
CC NICOTINAMIDE + (ADP-D-RIBOSYL)[N+1]-ACCEPTOR.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLE OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,  
CC AND BRAIN. LOW IN LIVER.  
CC -!- THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR  
CC CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-  
CC RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL  
CC ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN  
CC LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
-----  
This swiss-prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its



CC -1- COFACTOR: ZINC, CONTAINS TWO MOLE OF ZINC PER MOLE OF PROTEIN.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR  
CC CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-  
CC RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL  
CC ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN  
CC LENGTH OF 20-30 UNITS.  
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U94340; G2896792; -.  
DR EMBL; X65496; G56848; -.  
DR EMBL; X65497; G56850; ALT INIT.  
DR PROSITE; PS00347; PARP\_2N\_FINGER\_1; 2.  
DR PROSITE; PS50064; PARP\_2N\_FINGER\_2; 2.  
DR PFAM; PF00644; PARP; 1.  
DR HSP; P26446; PARP; 1.  
DR TRANSFERASE; GLYCOSYLTRANSFERASE; NAD; DNA-BINDING; NUCLEAR PROTEIN;  
KW ADP-RIBOSYLATION; ZINC-FINGER; ZINC.  
FT INIT\_MET 0 BY SIMILARITY.  
FT MOD\_RES 1 1  
FT DNA\_BIND 1 372  
FT DOMAIN 373 523 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 524 1013 NAD-BINDING.  
FT ZN\_FING 20 55 PARP-TYPE.  
FT ZN\_FING 124 161 PARP-TYPE.  
FT DOMAIN 206 208 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT MOD\_RES 407 407 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 413 413 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 435 435 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 437 437 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 444 444 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 456 456 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 484 484 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 488 488 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 491 491 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 512 512 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 513 513 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 519 519 ADP-RIBOSYL[N] (POTENTIAL).  
FT ACT\_SITE 892 892 BY SIMILARITY.  
FT CONFLICT 638 638 Y -> H (IN REF. 4).  
FT CONFLICT 641 641 E -> A (IN REF. 4).  
FT CONFLICT 752 752 N -> D (IN REF. 4).  
SQ SEQUENCE 1013 AA; 112529 MW; E5659A59 CRC32;  
  
Query Match 50.8%; Score 61; DB 1; Length 1013;  
Best Local Similarity 47.4%; Pred. No. 9.57e-01;  
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
Db 187 SAEDKEALKKOLPAVKSEG 205  
1:|||||  
QY 2 SGSDROTAKRLPAASGG 20  
-----  
RESULT 4  
ID HM09\_CAEEL STANDARD; PRT; 60 AA.  
AC P56407;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE HOMEBOX PROTEIN CEH-9 (FRAGMENT).  
GN CEH-9.  
OS CAENORHABDITIS ELEGANS.  
CC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
FT

OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91045075.  
RT HAWKINS N.C., MCGHEE J.D.;  
RT "Homeobox containing genes in the nematode *Caenorhabditis elegans*.";  
RL NUCLEIC ACIDS RES. 18:6101-6106(1990).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X52811; -; NOT\_ANNOTATED\_CDS.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS50071; HOMEBOX\_2; 1.  
DR HSP; P02833; ISAN.  
KW HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN.  
FT NON\_TER 1 1  
FT DNA\_BIND 1 60 HOMEBOX.  
FT NON\_TER 60 60  
FT SEQUENCE 60 AA; 7308 MW; 5CE6CD39 CRC32;  
  
Query Match 50.0%; Score 60; DB 1; Length 60;  
Best Local Similarity 52.9%; Pred. No. 1.50e+00;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
Db 27 SSSDRSELAKRLDVTER 43  
1:|||||  
QY 2 SGSDROTAKRLPAAS 18  
-----  
RESULT 5  
ID ATCL\_SYN7 STANDARD; PRT; 926 AA.  
AC P37278;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE CATION-TRANSPORTING ATPASE PAEL (EC 3.6.1.-).  
GN PAEL.  
OS SYNECHOCOCCUS SP. (STRAIN PCC 7942) (ANACYSTIS NIDULANS R2).  
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 93380581.  
RA KANAMARU K., KASHIWAGI S., MIZUNO T.;  
RT "The cyanobacterium, *Synechococcus* sp. PCC7942, possesses two  
RT distinct genes encoding cation-transporting p-type ATPases.";  
RL FEBS LETT. 330:99-104(1993).  
CC -1- FUNCTION: COULD POSSIBLY BE A CALCIUM ATPASE.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
CC (E1-E2 ATPASES).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D16436; G435123; -.  
DR PIR; S36742; S36742.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
DR PFAM; PF00122; E1-E2\_ATPase; 1.  
KW HYDROLASE; TRANSMEMBRANE; PHOSPHORYLATION; MAGNESIUM; ATP-BINDING.  
FT TRANSMEM 72 90 POTENTIAL.



GN CEH-23 OR ZK652.5.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE; 94130718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FRASER A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,  
RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,  
RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISSTER N.,  
RA LATREILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,  
RA SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,  
RA WATSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,  
RA WOLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL NATURE 368:32-38(1994).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE DISTAL-LESS FAMILY OF HOMEBOX  
CC PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L14429; G289773; -.  
DR PIR; S44907; S44907.  
DR WORMPEP; ZK652.5; CE00451.  
DR PROSITE; PS00027; HOMEBOX\_1; 1.  
DR PROSITE; PS00071; HOMEBOX\_2; 1.  
DR PFAM; PF00046; homeobox; 1.  
DR HSP; P22808; 1VND.  
KW HOMEBOX; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN.  
FT DNA\_BIND 211 270 HOMEBOX.  
SQ SEQUENCE 305 AA; 33260 MW; 7B4CC671 CRC32;  
Query Match 49.2%; Score 59; DB 1; Length 305;  
Best Local Similarity 43.8%; Pred. No. 2.33e+00;  
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
Db 238 GAERENLAORGLSPS 253  
::: || || ::||  
Qy 3 GSDRDTLAKRLPAAS 18  
RESULT 8  
ID CH62\_STRAL STANDARD; PRT; 539 AA.  
AC Q00768;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE 60 KD CHAPERONIN 2 (PROTEIN CPN60 2) (GROEL PROTEIN 2) (HSP56).  
GN GROEL2.  
OS STREPTOMYCES ALBUS G.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIADAE;  
OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92041639.  
RA MAZODIER P., GUGLIELMI G., DAVIES J., THOMPSON C.J.;  
RT "Characterization of the groEL-like genes in Streptomyces albus.";  
RL J. BACTERIOL. 173:7382-7386(1991).  
RN [2]  
RP SEQUENCE OF 1-10.

RX MEDLINE; 92041638.  
RA GUGLIELMI G., MAZODIER P., THOMPSON C.J., DAVIES J.;  
RT "A survey of the heat shock response in four Streptomyces species  
RT reveals two groEL-like genes and three groEL-like proteins in  
RT Streptomyces albus.";  
RL J. BACTERIOL. 173:7374-7381(1991).  
CC -!- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND  
CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS  
CC CONDITIONS (BY SIMILARITY).  
CC -!- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF  
CC 7 SUBUNITS (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M7658; G153294; -.  
DR PIR; C41325; C41325.  
DR PROSITE; PS00296; CHAPERONINS\_CPN60; 1.  
DR PFAM; PF00118; Cpn60\_TCP1; 1.  
DR HSP; P06139; 1GRL.  
KW CHAPERONE; ATP-BINDING; MULTIGENE FAMILY.  
FT INIT\_MET 0 0  
SQ SEQUENCE 539 AA; 56586 MW; 6D94EA2F CRC32;  
Query Match 49.2%; Score 59; DB 1; Length 539;  
Best Local Similarity 42.1%; Pred. No. 2.33e+00;  
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
Db 354 DSDYDREKLQERLANVAGG 372  
: || ||: ||: ||:  
Qy 1 NSGSDRDTLAKRLPAASG 19  
RESULT 9  
ID OTE\_DROME STANDARD; PRT; 406 AA.  
AC P20240;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)  
DE OTEFIN.  
GN OTE.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90243644.  
RA PADAN R., NAINUDEL-EPSTEYN S., GOITEIN R., FAINSOD A., GRUENBAUM Y.;  
RT "Isolation and characterization of the Drosophila nuclear envelope  
RT otefin cdna.";  
RL J. BIOL. CHEM. 265:7808-7813(1990).  
CC -!- FUNCTION: NUCLEAR ENVELOPE PROTEIN.  
CC -!- SUBCELLULAR LOCATION: INNER NUCLEAR MEMBRANE-ASSOCIATED. IT IS  
CC A COMPONENT OF THE SPINDLE ENVELOPE DURING EARLY MITOTIC CYCLES.  
CC -!- DEVELOPMENTAL STAGE: THROUGHOUT THE DEVELOPMENTAL STAGES OF THE  
CC FRUIT FLY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X17495; G8314; -.

```
DR PIR: A35360; A35360.
DR FLYBASE; FBgn003022; Ote.
KW NUCLEAR PROTEIN; TRANSMEMBRANE.
FT DOMAIN 74 81 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 133 137 SER/THR-RICH.
FT DOMAIN 210 215 SER/THR-RICH.
FT DOMAIN 250 254 SER/THR-RICH.
FT DOMAIN 289 293 SER/THR-RICH.
FT TRANSMEM 389 406 POTENTIAL.
SQ SEQUENCE 406 AA; 44658 MW; 525158AF CRC32;

Query Match 48.3%; Score 58; DB 1; Length 406;
Best Local Similarity 44.4%; Pred. No. 3.60e+00;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Db 12 TDSRKVLKRLRASIG 29
QY 2 SGSDRTLAKRLPAASG 19
: | | | | | | | | | |

RESULT 10
ID RL27_BORBU STANDARD; PRT; 81 AA.
AC OS1721;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE 50S RIBOSOMAL PROTEIN L27.
GN RMA OR BB0780.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA LATHIGRA R., WHITE O., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., HICKEY E.K., GWINN M.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., RICHARDSON D.,
RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WAITHEY L., MCDONALD L., ARIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi."
RL NATURE 390:580-586(1997).
CC -1- SIMILARITY: BELONGS TO THE L27P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001177; G2688720; -.
DR PROSITE; PS00831; RIBOSOMAL_L27; 1.
DR TIGR; BB0780; -.
KW RIBOSOMAL PROTEIN.
SQ SEQUENCE 81 AA; 8748 MW; 45875878 CRC32;

Query Match 47.5%; Score 57; DB 1; Length 81;
Best Local Similarity 45.0%; Pred. No. 5.53e+00;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 9 SSKNGRDSIKRLGVKRSRG 28
QY 1 NSGSDRTLAKRLPAASG 20
: | | | | | | | | | |

RESULT 11
ID Y065_MYCGE STANDARD; PRT; 466 AA.
AC P47311; Q49281;
```

```
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN MG065.
GN MG065.
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE; 96026346.
RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
RA FLEISCHMANN R.D., BUT C.J., KERLAVAGE A.R., SUTTON G., KELLEY J.M.,
RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMANN J.L.,
RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
RA TOMB J.-F., DOUGHERTY B.A., BOTT K.F., HU P.-C., LUCIER T.S.,
RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL SCIENCE 270:397-403(1995).
RN [2]
RP SEQUENCE OF 393-466 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE; 94075230.
RA PETERSON S.N., HU P.-C., BOTT K.F., HUTCHISON C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing."
RL J. BACTERIOL. 175:7918-7930(1993).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U39686; G1045740; -.
DR EMBL; U02154; G406365; -.
DR TIGR; MG065; -.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
KW HYPOTHETICAL PROTEIN; ATP-BINDING; TRANSPORT.
FT NP BIND 269 276 ATP (POTENTIAL).
FT CONFLICT 394 399 GDEPTG -> WWTYV (IN REF. 2).
SQ SEQUENCE 466 AA; 54009 MW; 4539B495 CRC32;

Query Match 47.5%; Score 57; DB 1; Length 466;
Best Local Similarity 57.1%; Pred. No. 5.53e+00;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 361 DSLQKKLPNELSG 374
QY 7 DTLAKRLPAASG 20
: | | | | | | | | | |

RESULT 12
ID FPOL_MOUSE STANDARD; PRT; 1012 AA.
AC P11103;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE POLY [ADP-RIBOSE] POLYMERASE (EC 2.4.2.30) (PARE) (ADPRT) (NAD(+)) ADP-
DE RIBOSYLTRANSFERASE) (POLY[ADP-RIBOSE] SYNTHETASE).
GN ADPRT OR ADPRT.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BXSB;
```



```
CC CC -!- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC CC AMINOTRANSFERASES.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: U15034; G693692; -
CC CC PROSITE: PS00105; AA_TRANSFER_CLASS_1; 1.
CC CC PFAM: PF00155; aminotran_1; 1.
CC CC HSSP: P00508; ITAT.
CC CC KW TRANSFERASE; AMINOTRANSFERASE; PYRIDOXAL PHOSPHATE; CHLOROPLAST;
CC CC TRANSIT PEPTIDE; MULTIGENE FAMILY.
CC CC FT TRANSIT 1 43 CHLOROPLAST (POTENTIAL).
CC CC FT CHAIN 44 449 ASPARTATE AMINOTRANSFERASE.
CC CC FT BINDING 295 295 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC CC SQ SEQUENCE 449 AA; 48954 MW; 9354CEC3 CRC32;
CC CC -----
CC CC Query Match 46.7%; Score 56; DB 1; Length 449;
CC CC Best Local Similarity 50.0%; Pred.No. 8.45e+00;
CC CC Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
CC CC -----
CC CC Db 24 NSGSDSDNLSLVSPTSG 43
CC CC ||||| I I : : : : |||
CC CC QY 1 NSGSDRDTLAKRLPAAASGG 20
CC CC -----
CC CC RESULT 15
CC CC ID ACHD_BOVIN STANDARD; PRT: 516 AA.
CC CC AC P04759;
CC CC DT 13-AUG-1987 (REL. 05, CREATED)
CC CC DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
CC CC DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
CC CC DE ACETYLCHOLINE RECEPTOR PROTEIN, DELTA CHAIN PRECURSOR.
CC CC GN CHRD.
CC CC OS BOS TAURUS (BOVINE).
CC CC OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC CC OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
CC CC [1]
CC CC RN SEQUENCE FROM N.A.
CC CC RX MEDLINE: 85203892.
CC CC RA KUBO S., NODA M., TAKAI T., TANABE T., KAYANO T., SHIMIZU S.,
CC CC RA TANAKA K., TAKAHASHI H., HIROSE T., INAYAMA S., KIKUNO R., MIYATA T.,
CC CC RA NUMA S.;
CC CC RT "Primary structure of delta subunit precursor of calf muscle
CC CC acetylcholine receptor deduced from cDNA sequence.";
CC CC RL EUR. J. BIOCHEM. 149:5-13(1985).
CC CC CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
CC CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
CC CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
CC CC MEMBRANE.
CC CC CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
CC CC DELTA, AND GAMMA (IN IMMATURE MUSCLE) OR EPSILON (IN MATURE
CC CC MUSCLE) CHAINS.
CC CC CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC CC CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: X02473; G757751; -
CC CC PIR: A24613; A24613.
CC CC PROSITE: PS00236; NEUROTRF_ION_CHANNEL; 1.
CC CC PFAM: PF00065; neur_chan; 1.
```

```
KW RECEPTOR; POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; SIGNAL;
KW TRANSMEMBRANE; PHOSPHORYLATION.
FT SIGNAL 1 21
FT CHAIN 22 516 ACETYLCHOLINE RECEPTOR PROTEIN, DELTA.
FT DOMAIN 22 244 EXTRACELLULAR.
FT TRANSMEM 245 269
FT TRANSMEM 279 296
FT TRANSMEM 311 332
FT DOMAIN 333 470
FT TRANSMEM 471 493 CYTOPLASMIC.
FT DISULFID 150 164 BY SIMILARITY.
FT CARBOHYD 96 96 POTENTIAL.
FT CARBOHYD 163 163 PROBABLE.
FT MOD_RES 389 389 PHOSPHORYLATION (BY TYR-KINASES)
FT (BY SIMILARITY).
SQ SEQUENCE 516 AA; 58587 MW; 1E04EF07 CRC32;
Query Match 46.7%; Score 56; DB 1; Length 516;
Best Local Similarity 53.8%; Pred.No. 8.45e+00;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 404 SERHGLARLTAA 416
QY 4 SDRDTLAKRLPAA 16
I : I || : || : I
```

Search completed: Mon Aug 23 13:07:42 1999  
Job time : 10 secs.





RA	ROUNSELEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C., STOKES S.M., MASON T.M., KERLAVAGE A.R., ADAMS M.D., SOMERVILLE C.R., VENTER J.C.;
RA	SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR	EMBL; AC002332; G2459442; -
KW	DNA-BINDING.
SQ	SEQUENCE 296 AA; 30296 MW; 8059D402 CRC32;
	Query Match 49.2%; Score 59; DB 10; Length 296;
	Best Local Similarity 40.0%; Pred. No. 4.72e+00;
	Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
Db	13 NSAGENSEVLNNLPGGESGG 32
Qy	1 NSGSDRDTLAKRLPAAASGG 20
	:  ::
RESULT 5	
ID O23142	PRELIMINARY; PRT; 351 AA.
AC O23142;	
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)	
DE AT-HOOK PROTEIN 1.	
DE AHPI.	
GN AHPI.	
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).	
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; TRACHEOPHYTA;	
OC EUPHYLLIPHTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;	
OC CAPPALES; BRASSICACEAE; ARABIDOPSIS.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=FLOWER BUDS;	
RC HOFMANN W.A., SAEDLER H., HUIJSER P.;	
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
DR EMBL; A222585; E1172826; -	
SQ SEQUENCE 351 AA; 36317 MW; DDB8DC14 CRC32;	
	Query Match 49.2%; Score 59; DB 10; Length 351;
	Best Local Similarity 40.0%; Pred. No. 4.72e+00;
	Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
Db	68 NSAGENSEVLNNLPGGESGG 87
Qy	1 NSGSDRDTLAKRLPAAASGG 20
	:  ::
RESULT 6	
ID O33658	PRELIMINARY; PRT; 477 AA.
AC O33658;	
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)	
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)	
DE 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN)	
DE (HEAT SHOCK PROTEIN 60).	
GN GROEL2.	
OS STREPTOMYCES LIVIDANS.	
OC BACTERIA; FRMICUTES; ACTINOBACTERIA; ACTINOBACTERIADAE;	
OC ACTINOMYCETALES; STREPTOMYCINEAE; STREPTOMYCETACEAE; STREPTOMYCES.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=Tk21;	
RC MEDLINE; 98048481.	
RA DE LEON P., MARCO S., ISIEGAS C., MARINA A., CARRASCOSA J.L., MELLADO R.P.;	
RA lividans groes, groel1 and groel2 genes.";	
RL Microbiology 143:3563-3571(1997).	
-1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS (BY SIMILARITY).	
-1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF 7 SUBUNITS (BY SIMILARITY).	
-1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.	
DR EMBL; X95971; E226026; -	

DR PROSITE: PS00296: CHAPERONINS\_CPN60; 1.  
DR PFAM: PF00118: CPN60\_TCPI; 1.  
KW CHAPERONE; ATP-BINDING.  
SQ SEQUENCE 477 AA; 50660 MW; 31EFAEB CRC32;

Query Match 49.28; Score 59; DB 2; Length 477;  
Best Local Similarity 42.18; Pred. No. 4.72e+00;  
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 355 DSDYDREKLOERLANVAGG 373  
QY 1 NSGSDRDTLAKRLPAAAS 19  
| | | | | | | | | |

RESULT 7 PRELIMINARY; PRT; 358 AA.  
ID O62895  
AC O62895;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DE MHC CLASS I ALPHA CHAIN.  
GN ICPU-F9.  
OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;  
OC ICTALURUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANTAO A.B., CHINCHAR V.G., MCCONNELL T.J., MILLER N.W., CLEM L.W.,  
RA WILSON M.R.;  
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF053548; G2995929; -.  
KW MHC.  
SQ SEQUENCE 358 AA; 39855 MW; F7018615 CRC32;

Query Match 48.38; Score 58; DB 7; Length 358;  
Best Local Similarity 43.88; Pred. No. 7.18e+00;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 195 GYGRETLEKRVPTAS 210  
QY 3 GSDRDTLAKRLPAAAS 18  
| | | | | | | | | |

RESULT 8 PRELIMINARY; PRT; 358 AA.  
ID O62894  
AC O62894;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DE MHC CLASS I ALPHA CHAIN.  
GN ICPU-F4.  
OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;  
OC ICTALURUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANTAO A.B., CHINCHAR V.G., MCCONNELL T.J., MILLER N.W., CLEM L.W.,  
RA WILSON M.R.;  
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF053547; G2995927; -.  
KW MHC.  
SQ SEQUENCE 358 AA; 39885 MW; 77A558B3 CRC32;

Query Match 48.38; Score 58; DB 7; Length 358;  
Best Local Similarity 43.88; Pred. No. 7.18e+00;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 195 GYGRETLEKRVPTAS 210  
QY 3 GSDRDTLAKRLPAAAS 18  
| | | | | | | | | |

RESULT 9 PRELIMINARY; PRT; 361 AA.  
ID O62896  
AC O62896;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DE MHC CLASS I ALPHA CHAIN.  
GN ICPU-E7.  
OS ICTALURUS PUNCTATUS (CHANNEL CATFISH).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;  
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; SILURIFORMES; ICTALURIDAE;  
OC ICTALURUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ANTAO A.B., CHINCHAR V.G., MCCONNELL T.J., MILLER N.W., CLEM L.W.,  
RA WILSON M.R.;  
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AF053549; G2995931; -.  
KW MHC.  
SQ SEQUENCE 361 AA; 40914 MW; 134453BE CRC32;

Query Match 48.38; Score 58; DB 7; Length 361;  
Best Local Similarity 43.88; Pred. No. 7.18e+00;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 195 GYGRETLEKRVPTAS 210  
QY 3 GSDRDTLAKRLPAAAS 18  
| | | | | | | | | |

RESULT 10 PRELIMINARY; PRT; 519 AA.  
ID Q18671  
AC Q18671;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE C47D12.8 PROTEIN (FRAGMENT).  
GN C47D12.8.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GAJADSTY S.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT \*2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans\*;  
RL NATURE 368:32-38(1994).  
DR EMBL; 269902; E1344710; -.  
FT NON\_TER 519 519  
SQ SEQUENCE 519 AA; 59679 MW; 98B051E4 CRC32;

Query Match 48.38; Score 58; DB 5; Length 519;  
Best Local Similarity 66.78; Pred. No. 7.18e-00;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 28 ERATLAKTLPAS 39  
QY 5 DRTLAKRLPAA 16  
| | | | | | | | | |

```

RESULT 11
ID O28948          PRELIMINARY;      PRT;    547 AA.
AC O28948;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE PHENYLALANYL-TRNA SYNTHETASE, SUBUNIT BETA (PHET).
GN AF1424.
OS ARCHAEoglobus fulgidus.
OC ARCHAEA: EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobaceae;
OC ARCHAEoglobus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE; 98049343.
RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KIRPIDES N.C.,
RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
RA KIRNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,
RA COTTON M.D., SPRIGGS T., ARTIACH P., KATNE B.P., SYKES S.M.,
RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA VENTER J.C.;
RT "The complete genome sequence of the hyperthermophilic,
RT sulphate-reducing archaeon Archaeoglobus fulgidus.";
RL NATURE 390:364-370(1997).
DR EMBL; AE001005; G2649146; -.
DR TIGR; AF1424; -.
KW HYPOTHETICAL PROTEIN; AMINOACYL-TRNA SYNTHETASE.
SQ SEQUENCE 547 AA; 62278 MW; 4AC4E641 CRC32;

Query Match      48.3%; Score 58; DB 1; Length 547;
Best Local Similarity 56.7%; Pred. No. 7.18e+00;
Matches      8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db      16 GADRTILSRLP 27
Qy      3 GSDRTLAKRLP 14
      1:|||||: |||
      3 GSDRTLAKRLP 14

RESULT 12
ID O87839          PRELIMINARY;      PRT;    560 AA.
AC O87839;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PUTATIVE TPP-REQUIRING ENZYME.
GN SC8A6.05C.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA: FIRMICUTES; ACTINOBACTERIA: ACTINOBACTERIDAE;
OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA SEEGER K.J., HARRIS D.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmid and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. MICROBIOL. 21:77-96(1996).
CC -!- SIMILARITY: WITH OTHER ENZYMES WHICH REQUIRE TPP.
DR EMBL; AL031013; E1309411; -.
KW PROSITE; PS00187; TPP_ENZYMES; 1.
KW FLAVOPROTEIN; LVASE; THIAMINE PYROPHOSPHATE.
SQ SEQUENCE 560 AA; 60104 MW; 24AFB209 CRC32;

Query Match      48.3%; Score 58; DB 2; Length 560;
Best Local Similarity 50.0%; Pred. No. 7.18e+00;
Matches      6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db      351 AAERTALDKRLP 362
Qy      3 GSDRTLAKRLP 14
      ::::|:| ||||
      3 GSDRTLAKRLP 14

RESULT 13
ID Q53039          PRELIMINARY;      PRT;    104 AA.
AC Q53039;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE NITRILE HYDRATASE-B HOMOLOGUE.
GN NHG.
OS RHODOCOCOCCUS RHODOCHROUS.
OC BACTERIA: FIRMICUTES; ACTINOBACTERIA: ACTINOBACTERIDAE;
OC ACTINOMYCETALES; CORYNEBACTERIINAE; NOCARDIACEAE; RHODOCOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J1;
RX MEDLINE; 96210630.
RA KOMEDA H., KOBAYASHI M., SHIMIZU S.;
RT "Characterization of the gene cluster of high-molecular-mass nitrile
RT hydratase (H-NHase) induced by its reaction product in Rhodococcus
RT rhodochrous J1.";
RL PROC. NATL. ACAD. SCI. U.S.A. 93:4267-4272(1996).
DR EMBL; D67027; D1011705; ...
SQ SEQUENCE 104 AA; 11650 MW; 5C3C2153 CRC32;

Query Match      46.7%; Score 56; DB 2; Length 104;
Best Local Similarity 50.0%; Pred. No. 1.63e+01;
Matches      7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db      4 DTLDRLPATGTAA 17
Qy      7 DTLAKRLPAAASGG 20
      |||: |||||:::
      7 DTLAKRLPAAASGG 20

RESULT 14
ID O54166          PRELIMINARY;      PRT;    122 AA.
AC O54166;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 12.7 KD PROTEIN.
GN SC7H1.07C.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA: FIRMICUTES; ACTINOBACTERIA: ACTINOBACTERIDAE;
OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MURPHY L., HARRIS D.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
```

RX MEDLINE: 97000351.  
RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,  
RA KINASHI H., HOPWOOD D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL MOL. MICROBIOL. 21:77-96(1996).  
DR EMBL: AL021411; E1245736;  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 122 AA; 12657 MW; C611FB76 CRC32;

Query Match 46.7%; Score 56; DB 2; Length 122;  
Best Local Similarity 44.4%; Pred. No. 1.63e+01;  
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 103 GCGREPMSRTAAASAG 120  
QY 3 GSDRDTLAKRLPAAASGG 20  
| : : : | : ||| : |  
| : : : | : ||| : |

RESULT 15  
ID C33037 PRELIMINARY; PRT; 302 AA.  
AC C33037;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE XERC PROTEIN.  
DE XERC.  
GN XERC.  
OS MYCOBACTERIUM LEPRAE.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;  
OC ACTINOMYCETALES; CORYNEBACTERINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA SEGER K.J., HARRIS D.;  
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93188700.  
RA EIGLMEIER K., HONORE N., WOODS S.A., CAUDRON B., COLE S.T.;  
RT "Use of an ordered cosmid library to deduce the genomic organization  
RT of Mycobacterium leprae.";  
RL MOL. MICROBIOL. 7:197-206(1993).  
DR EMBL: Z97369; E1192373;  
DR PFAM: PF00589; Phage\_integrase; 1.  
SQ SEQUENCE 302 AA; 32705 MW; 76D7B053 CRC32;

Query Match 46.7%; Score 56; DB 2; Length 302;  
Best Local Similarity 50.0%; Pred. No. 1.63e+01;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 66 AGAARTTLARRISA 79  
QY 2 GSDRDTLAKRLPA 15  
| : : : | : ||| : |  
| : : : | : ||| : |

Search completed: Mon Aug 23 13:08:28 1999  
Job time : 27 secs.

**This Page Blank (uspto)**

(TM)

Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match						
1	6	25.0	30	159	30	W52839	Secreted protein enco	2.33e+01
2	6	25.0	31	35	35	W75916	Human p38 receptor pr	2.33e+01
3	5	20.8	17	14	14	R76097	Omega conotoxin GVIA-	2.49e+02
4	5	20.8	17	16	16	R32785	GVIA omega conotoxin	2.49e+02
5	5	20.8	27	39	39	W95566	Omega-conopeptide GVI	2.49e+02
6	5	20.8	27	37	37	R72607	Conus genus natural o	2.49e+02
7	5	20.8	27	10	10	W51035	N-type calcium channe	2.49e+02
8	5	20.8	27	20	20	W12969	Omega conopeptide SNX	2.49e+02
9	5	20.8	27	23	23	W19546	Natural omega-conopep	2.49e+02
10	5	20.8	27	7	7	R37754	GVIA/SNX-124.	2.49e+02
11	5	20.8	27	8	8	R39610	GVIA/SNX124.	2.49e+02
12	5	20.8	27	14	14	R76091	Omega conotoxin GVIA	2.49e+02
13	5	20.8	27	6	6	R32779	GVIA omega conotoxin	2.49e+02
14	5	20.8	47	27	27	W32240	B chain sequence of a	2.49e+02
15	5	20.8	47	17	17	R88034	B chain sequence of m	2.49e+02
16	5	20.8	50	17	17	R93032	Monellin B chain.	2.49e+02

90	5	20.8	141 18	W00105	Generic ob protein, I	2.49e+02	163	5	20.8	146 30	W50999	Anti-obesity protein	2.49e+02
91	5	20.8	141 18	W00050	N-terminally truncate	2.49e+02	164	5	20.8	146 30	W51000	Anti-obesity protein	2.49e+02
92	5	20.8	144 20	W03524	Anti-obesity protein.	2.49e+02	165	5	20.8	146 34	W57818	Anti-obesity protein.	2.49e+02
93	5	20.8	144 18	W00049	N-terminally truncate	2.49e+02	166	5	20.8	146 34	W57816	Anti-obesity protein.	2.49e+02
94	5	20.8	144 20	W03523	Generic anti-obesity	2.49e+02	167	5	20.8	146 35	W59920	Leptin receptor agon	2.49e+02
95	5	20.8	145 18	W00302	Human delta Gln28 ob	2.49e+02	168	5	20.8	146 20	W06524	Wild type ob protein.	2.49e+02
96	5	20.8	145 23	W23326	Obesity protein analo	2.49e+02	169	5	20.8	146 26	W34482	Human obesity protein	2.49e+02
97	5	20.8	145 17	W00541	Human mature obesity	2.49e+02	170	5	20.8	146 36	W18833	Obesity protein analo	2.49e+02
98	5	20.8	145 27	W30893	Synthetic obesity pro	2.49e+02	171	5	20.8	146 36	W18837	Obesity protein analo	2.49e+02
99	5	20.8	146 38	W80500	Leptin receptor analo	2.49e+02	172	5	20.8	146 36	W18836	Obesity protein analo	2.49e+02
100	5	20.8	146 38	W80501	Leptin receptor analo	2.49e+02	173	5	20.8	146 36	W85595	Properly folded obesi	2.49e+02
101	5	20.8	146 38	W80502	Leptin receptor analo	2.49e+02	174	5	20.8	146 36	W1307	Obesity protein analo	2.49e+02
102	5	20.8	146 36	W69687	Obesity protein analo	2.49e+02	175	5	20.8	146 23	W10153	Properly folded obesi	2.49e+02
103	5	20.8	146 36	W10301	Obesity protein analo	2.49e+02	176	5	20.8	146 23	W22903	Biologically active o	2.49e+02
104	5	20.8	146 36	W17302	Obesity protein analo	2.49e+02	177	5	20.8	146 26	W34487	Obesity protein analo	2.49e+02
105	5	20.8	146 36	W80499	Leptin receptor analo	2.49e+02	178	5	20.8	146 23	W11279	Generic leptin mimeti	2.49e+02
106	5	20.8	146 18	R9494	Chimeric ob protein (	2.49e+02	179	5	20.8	146 28	W45472	Anti-obesity protein,	2.49e+02
107	5	20.8	146 36	W17299	Obesity protein analo	2.49e+02	180	5	20.8	146 28	W45473	Anti-obesity protein,	2.49e+02
108	5	20.8	146 26	W34489	Obesity protein analo	2.49e+02	181	5	20.8	146 36	W1840	Obesity protein analo	2.49e+02
109	5	20.8	146 33	W63606	Preferred obesity (Ob	2.49e+02	182	5	20.8	146 23	W22896	Biologically active o	2.49e+02
110	5	20.8	146 25	W32573	Anti obesity protein	2.49e+02	183	5	20.8	146 30	W53334	Obesity protein analo	2.49e+02
111	5	20.8	146 32	W62306	Anti-obesity protein	2.49e+02	184	5	20.8	146 30	W53333	Obesity protein analo	2.49e+02
112	5	20.8	146 32	W62310	Anti-obesity protein	2.49e+02	185	5	20.8	146 36	W18843	Obesity protein analo	2.49e+02
113	5	20.8	146 36	W17311	Obesity protein analo	2.49e+02	186	5	20.8	146 36	W1845	Obesity protein analo	2.49e+02
114	5	20.8	146 36	W17310	Obesity protein analo	2.49e+02	187	5	20.8	146 36	W18835	Obesity protein analo	2.49e+02
115	5	20.8	146 35	W69688	Obesity protein analo	2.49e+02	188	5	20.8	146 36	W1884	Obesity protein analo	2.49e+02
116	5	20.8	146 35	W69695	Obesity protein analo	2.49e+02	189	5	20.8	146 30	W53340	Obesity protein analo	2.49e+02
117	5	20.8	146 36	W17304	Obesity protein analo	2.49e+02	190	5	20.8	146 30	W53339	Obesity protein analo	2.49e+02
118	5	20.8	146 36	W17303	Obesity protein analo	2.49e+02	191	5	20.8	146 30	W53337	Obesity protein analo	2.49e+02
119	5	20.8	146 36	W17309	Obesity protein analo	2.49e+02	192	5	20.8	146 30	W53338	Obesity protein analo	2.49e+02
120	5	20.8	146 36	W17308	Obesity protein analo	2.49e+02	193	5	20.8	146 23	W11275	Human leptin for trea	2.49e+02
121	5	20.8	146 28	W45474	Anti-obesity protein.	2.49e+02	194	5	20.8	146 23	W08596	Properly folded obesi	2.49e+02
122	5	20.8	146 28	W45478	Anti-obesity protein.	2.49e+02	195	5	20.8	146 31	W51005	Anti-obesity protein	2.49e+02
123	5	20.8	146 18	W00014	Chimeric ob protein (	2.49e+02	196	5	20.8	146 31	W51003	Anti-obesity protein	2.49e+02
124	5	20.8	146 18	W00301	Human ob protein.	2.49e+02	197	5	20.8	146 18	W00033	Acid stable modified	2.49e+02
125	5	20.8	146 25	W32578	Anti obesity protein	2.49e+02	198	5	20.8	146 18	R99495	Chimeric ob protein (	2.49e+02
126	5	20.8	146 27	W30892	Synthetic obesity pro	2.49e+02	199	5	20.8	146 26	W34485	Obesity protein analo	2.49e+02
127	5	20.8	146 33	W18630	Obesity protein analo	2.49e+02	200	5	20.8	146 26	W34486	Obesity protein analo	2.49e+02
128	5	20.8	146 35	W26191	Obesity protein analo	2.49e+02	201	5	20.8	146 36	W18834	Obesity protein analo	2.49e+02
129	5	20.8	146 33	W62608	Preferred obesity (Ob	2.49e+02	202	5	20.8	146 36	W18838	Obesity protein analo	2.49e+02
130	5	20.8	146 23	W11272	Murine leptin for tre	2.49e+02	203	5	20.8	146 36	W18839	Obesity protein analo	2.49e+02
131	5	20.8	146 27	W30894	Synthetic obesity pro	2.49e+02	204	5	20.8	146 25	W26196	Obesity protein analo	2.49e+02
132	5	20.8	146 33	W62604	Formula IV for an obe	2.49e+02	205	5	20.8	146 25	W26195	Obesity protein analo	2.49e+02
133	5	20.8	146 33	W62607	Preferred obesity (Ob	2.49e+02	206	5	20.8	146 30	W53332	Obesity protein analo	2.49e+02
134	5	20.8	146 25	W26193	Obesity protein analo	2.49e+02	207	5	20.8	146 30	W53342	Obesity protein analo	2.49e+02
135	5	20.8	146 23	W11278	Generic leptin mimeti	2.49e+02	208	5	20.8	146 35	W69682	Human obesity protein	2.49e+02
136	5	20.8	146 23	W18629	Obesity protein analo	2.49e+02	209	5	20.8	146 35	W69683	Leptin receptor agon	2.49e+02
137	5	20.8	146 35	W69692	Obesity protein analo	2.49e+02	210	5	20.8	146 35	W59919	Leptin receptor agon	2.49e+02
138	5	20.8	146 35	W69693	Obesity protein analo	2.49e+02	211	5	20.8	146 31	W50994	Anti-obesity protein	2.49e+02
139	5	20.8	146 36	W17305	Obesity protein analo	2.49e+02	212	5	20.8	146 31	W50995	Anti-obesity protein	2.49e+02
140	5	20.8	146 36	W17306	Obesity protein analo	2.49e+02	213	5	20.8	146 18	R99493	Chimeric ob protein (	2.49e+02
141	5	20.8	146 25	W26194	Obesity protein analo	2.49e+02	214	5	20.8	146 35	W70369	Leptin receptor ligan	2.49e+02
142	5	20.8	146 18	W00048	Generic mammalian Ob	2.49e+02	215	5	20.8	146 31	W50998	Anti-obesity protein	2.49e+02
143	5	20.8	146 36	W00015	Chimeric ob protein (	2.49e+02	216	5	20.8	146 31	W50993	Anti-obesity protein	2.49e+02
144	5	20.8	146 36	W17842	Obesity protein analo	2.49e+02	217	5	20.8	146 26	W34495	Obesity protein analo	2.49e+02
145	5	20.8	146 36	W17841	Obesity protein analo	2.49e+02	218	5	20.8	146 31	W50996	Anti-obesity protein	2.49e+02
146	5	20.8	146 35	W69684	Obesity protein analo	2.49e+02	219	5	20.8	146 31	W51004	Anti-obesity protein	2.49e+02
147	5	20.8	146 23	W18621	Obesity protein analo	2.49e+02	220	5	20.8	146 26	W34492	Obesity protein analo	2.49e+02
148	5	20.8	146 24	W24031	Human obesity protein	2.49e+02	221	5	20.8	146 35	W69689	Obesity protein analo	2.49e+02
149	5	20.8	146 35	W69690	Obesity protein analo	2.49e+02	222	5	20.8	146 35	W69685	Obesity protein analo	2.49e+02
150	5	20.8	146 35	W69691	Obesity protein analo	2.49e+02	223	5	20.8	146 35	W69686	Obesity protein analo	2.49e+02
151	5	20.8	146 25	W26200	Obesity protein analo	2.49e+02	224	5	20.8	146 35	W69694	Obesity protein analo	2.49e+02
152	5	20.8	146 23	W10151	Properly folded obesi	2.49e+02	225	5	20.8	146 32	W57258	Amino acid sequence o	2.49e+02
153	5	20.8	146 23	W22900	Biologically active o	2.49e+02	226	5	20.8	146 32	W57257	Amino acid sequence o	2.49e+02
154	5	20.8	146 25	W26192	Obesity protein analo	2.49e+02	227	5	20.8	146 27	W30900	Synthetic obesity pro	2.49e+02
155	5	20.8	146 18	W00011	Acid stable modified	2.49e+02	228	5	20.8	146 27	W36448	Synthetic obesity pro	2.49e+02
156	5	20.8	146 28	W45476	Anti-obesity protein,	2.49e+02	229	5	20.8	146 27	W22901	Biologically active o	2.49e+02
157	5	20.8	146 31	W51002	Anti-obesity protein	2.49e+02	230	5	20.8	146 23	W10152	Properly folded obesi	2.49e+02
158	5	20.8	146 31	W51001	Anti-obesity protein	2.49e+02	231	5	20.8	146 33	W62611	Preferred obesity (Ob	2.49e+02
159	5	20.8	146 25	W32579	Anti obesity protein	2.49e+02	232	5	20.8	146 23	W18625	Obesity protein analo	2.49e+02
160	5	20.8	146 25	W32580	Anti obesity protein	2.49e+02	233	5	20.8	146 27	W30898	Synthetic obesity pro	2.49e+02
161	5	20.8	146 28	W45470	Anti-obesity protein,	2.49e+02	234	5	20.8	146 26	W34488	Obesity protein analo	2.49e+02
162	5	20.8	146 28	W45475	Anti-obesity protein,	2.49e+02	235	5	20.8	146 18	R99490	Chimeric ob protein.	2.49e+02



236	5	20.8	146 36	W71298	Obesity protein analo	2.49e+02	309	146 33	W49019	Mutant human obese pr	2.49e+02
237	5	20.8	146 34	W57815	Anti-obesity protein.	2.49e+02	310	146 33	W49020	Mutant human obese pr	2.49e+02
238	5	20.8	146 35	W34496	Obesity protein analo	2.49e+02	311	146 18	R99500	Acid stable modified	2.49e+02
239	5	20.8	146 26	W34483	Human obesity protein	2.49e+02	312	146 18	W00012	Acid stable modified	2.49e+02
240	5	20.8	146 30	W53330	Obesity protein analo	2.49e+02	313	146 33	W62602	Formula II for an obe	2.49e+02
241	5	20.8	146 30	W53331	Obesity protein analo	2.49e+02	314	146 26	W34491	Obesity protein analo	2.49e+02
242	5	20.8	146 34	W57817	Anti-obesity protein.	2.49e+02	315	146 18	R99491	Chimeric ob protein (	2.49e+02
243	5	20.8	146 34	W57817	Anti-obesity protein.	2.49e+02	316	146 33	W62601	Formula I for an obe	2.49e+02
244	5	20.8	146 25	W26189	Obesity protein analo	2.49e+02	317	146 27	W30897	Synthetic obesity pro	2.49e+02
245	5	20.8	146 33	W18627	Obesity protein analo	2.49e+02	318	146 18	R99497	Generic ob protein.	2.49e+02
246	5	20.8	146 33	W49022	Mutant human obese pr	2.49e+02	319	146 23	W18626	Obesity protein analo	2.49e+02
247	5	20.8	146 25	W32569	Anti obesity protein	2.49e+02	320	146 22	W07191	Generic, biologically	2.49e+02
248	5	20.8	146 26	W34397	Human Met-OB protein	2.49e+02	321	146 25	W26190	Obesity protein analo	2.49e+02
249	5	20.8	146 23	W08599	Properly folded obesi	2.49e+02	322	147 24	W27167	Human recombinant OB	2.49e+02
250	5	20.8	146 23	W22897	Biologically active o	2.49e+02	323	147 26	W34394	Human Met-OB protein.	2.49e+02
251	5	20.8	146 25	W32576	Anti obesity protein	2.49e+02	324	147 30	W53328	Human obesity protein	2.49e+02
252	5	20.8	146 25	W32577	Anti obesity protein	2.49e+02	325	147 24	W27170	Human recombinant OB	2.49e+02
253	5	20.8	146 28	W45477	Anti-obesity protein	2.49e+02	326	148 36	W71848	Obesity protein analo	2.49e+02
254	5	20.8	146 18	R99492	Chimeric ob protein (	2.49e+02	327	148 36	W71849	Obesity protein analo	2.49e+02
255	5	20.8	146 25	W26197	Obesity protein analo	2.49e+02	328	148 32	W62312	Anti-obesity protein	2.49e+02
256	5	20.8	146 24	W18338	Chimpanzee leptin.	2.49e+02	329	148 32	W62313	Anti-obesity protein	2.49e+02
257	5	20.8	146 25	W26198	Obesity protein analo	2.49e+02	330	148 34	W57820	Anti-obesity protein.	2.49e+02
258	5	20.8	146 24	W13639	Gorilla leptin.	2.49e+02	331	148 36	W71847	Generic obesity prote	2.49e+02
259	5	20.8	146 23	W18623	Obesity protein analo	2.49e+02	332	148 34	W57821	Anti-obesity protein.	2.49e+02
260	5	20.8	146 17	W00539	Human mature obesity	2.49e+02	333	148 22	W07192	Human anti-obesity pr	2.49e+02
261	5	20.8	146 18	R99489	Generic ob protein.	2.49e+02	334	148 32	W62314	Anti-obesity protein	2.49e+02
262	5	20.8	146 18	R99498	Generic ob protein (T	2.49e+02	335	148 32	W62315	Human ob protein SQ	2.49e+02
263	5	20.8	146 33	W62609	Preferred obesity (Ob	2.49e+02	336	148 24	W28801	Human mast protein.	2.49e+02
264	5	20.8	146 27	W30899	Synthetic obesity pro	2.49e+02	337	148 18	R99496	Generic Met-X-ob prot	2.49e+02
265	5	20.8	146 27	W30895	Synthetic obesity pro	2.49e+02	338	148 32	W62311	Anti-obesity protein	2.49e+02
266	5	20.8	146 30	W53341	Obesity protein analo	2.49e+02	339	148 34	W57825	Anti-obesity protein.	2.49e+02
267	5	20.8	146 28	W45469	Anti-obesity protein,	2.49e+02	340	148 34	W57822	Anti-obesity protein.	2.49e+02
268	5	20.8	146 26	W34493	Obesity protein analo	2.49e+02	341	148 34	W57823	Anti-obesity protein.	2.49e+02
269	5	20.8	146 26	W30791	Obesity protein compl	2.49e+02	342	148 34	W57824	Anti-obesity protein.	2.49e+02
270	5	20.8	146 24	W13840	Orangutan leptin.	2.49e+02	343	156 37	W68542	Leech coagulation Fac	2.49e+02
271	5	20.8	146 18	W00010	Acid stable modified	2.49e+02	344	159 27	W44125	Streptococcus pneumon	2.49e+02
272	5	20.8	146 18	R99499	Acid stable modified	2.49e+02	345	165 17	W00537	Human obesity protein	2.49e+02
273	5	20.8	146 25	W32571	Anti obesity protein	2.49e+02	346	166 17	W00531	Human obesity protein	2.49e+02
274	5	20.8	146 23	W22902	Biologically active o	2.49e+02	347	166 17	W00535	Human obesity protein	2.49e+02
275	5	20.8	146 32	W62307	Anti-obesity protein	2.49e+02	348	166 17	W00525	Human obesity protein	2.49e+02
276	5	20.8	146 23	W18624	Obesity protein analo	2.49e+02	349	166 17	W00515	Human obesity protein	2.49e+02
277	5	20.8	146 27	W30896	Synthetic obesity pro	2.49e+02	350	166 17	W00530	Human obesity protein	2.49e+02
278	5	20.8	146 26	W34484	Obesity protein analo	2.49e+02	351	166 17	W00523	Human obesity protein	2.49e+02
279	5	20.8	146 25	W34490	Obesity protein analo	2.49e+02	352	166 17	W00533	Human obesity protein	2.49e+02
280	5	20.8	146 32	W62305	Anti-obesity protein	2.49e+02	353	166 17	W00532	Human obesity protein	2.49e+02
281	5	20.8	146 32	W62304	Anti-obesity protein	2.49e+02	354	167 17	W00521	Human obesity protein	2.49e+02
282	5	20.8	146 23	W18628	Obesity protein analo	2.49e+02	355	167 17	W00520	Human obesity protein	2.49e+02
283	5	20.8	146 23	W23325	Obesity protein analo	2.49e+02	356	167 17	W00517	Human obesity protein	2.49e+02
284	5	20.8	146 23	W18622	Obesity protein analo	2.49e+02	357	167 17	R92720	Obesity protein.	2.49e+02
285	5	20.8	146 32	W62308	Anti-obesity protein	2.49e+02	358	167 17	W00518	Human obesity protein	2.49e+02
286	5	20.8	146 26	W32494	Obesity protein analo	2.49e+02	359	167 22	W03694	Human obese (ob) prot	2.49e+02
287	5	20.8	146 24	W24028	Mouse obesity protein	2.49e+02	360	167 22	W03694	Human obese (ob) prot	2.49e+02
288	5	20.8	146 25	W26201	Obesity protein analo	2.49e+02	361	167 31	W57442	Human leptin sequence	2.49e+02
289	5	20.8	146 25	W26199	Obesity protein analo	2.49e+02	362	167 27	W34060	Human obese (ob) prot	2.49e+02
290	5	20.8	146 25	W26199	Obesity protein analo	2.49e+02	363	167 17	W00519	Human obesity protein	2.49e+02
291	5	20.8	146 33	W49021	Mutant human obese pr	2.49e+02	364	167 18	R99478	Human ob protein.	2.49e+02
292	5	20.8	146 32	W62303	Anti-obesity protein	2.49e+02	365	171 8	R41488	Mouabatin platelet agg	2.49e+02
293	5	20.8	146 34	W57819	Anti-obesity protein.	2.49e+02	366	197 39	W73592	Factor Xa inhibitor p	2.49e+02
294	5	20.8	146 30	W53329	Obesity protein analo	2.49e+02	367	199 38	W85595	Mouse CD69.	2.49e+02
295	5	20.8	146 18	W00016	Chimeric ob protein (	2.49e+02	368	199 10	R54660	Marine CD69.	2.49e+02
296	5	20.8	146 34	W57814	Anti-obesity protein.	2.49e+02	369	202 12	R62517	Deduced sequence of 1	2.49e+02
297	5	20.8	146 33	W62603	Formula III for an ob	2.49e+02	370	212 35	W79363	Staphylococcus aureus	2.49e+02
298	5	20.8	146 33	W62605	Preferred obesity (Ob	2.49e+02	371	227 1	P94159	Part of H-GR1B precu	2.49e+02
299	5	20.8	146 30	W53335	Obesity protein analo	2.49e+02	372	257 9	R47190	Etherase.	2.49e+02
300	5	20.8	146 30	W53336	Obesity protein analo	2.49e+02	373	287 7	R34030	Fc-alpha-R.	2.49e+02
301	5	20.8	146 25	W32374	Anti obesity protein	2.49e+02	374	287 11	R59320	Human Fc-alpha-R.	2.49e+02
302	5	20.8	146 36	W71846	Generic obesity prote	2.49e+02	375	313 24	W21856	Protein used in prepa	2.49e+02
303	5	20.8	146 28	W45471	Anti-obesity protein,	2.49e+02	376	330 2	R24238	Vif polypeptide of FI	2.49e+02
304	5	20.8	146 33	W62610	Preferred obesity (Ob	2.49e+02	377	335 5	R25063	Soluble human IL-5 re	2.49e+02
305	5	20.8	146 25	W32568	Anti obesity protein	2.49e+02	378	335 6	R33699	shIL-5R-alpha.	2.49e+02
306	5	20.8	146 25	W32575	Anti obesity protein	2.49e+02	379	339 22	W20847	H. pylori surface or	2.49e+02
307	5	20.8	146 25	W32572	Anti obesity protein	2.49e+02	380	350 31	W56265	Mouse recombinase muR	2.49e+02
308	5	20.8	146 25	W32570	Anti obesity protein	2.49e+02	381	350 31	W56264	Human recombinase hsr	2.49e+02

382	5	20.8	361 24	W26001	Pectinase.	2.49e+02	455	4	16.7	28	8	R42717	Murine TGase3 27 kD f	2.22e+03
383	5	20.8	374 37	W83963	Recombinant human met	2.49e+02	456	4	16.7	33	10	R51040	Antibacterial peptide	2.22e+03
384	5	20.8	374 35	W49076	Recombinant human Met	2.49e+02	457	4	16.7	33	8	R40062	Hib OMP P1 peptide HI	2.22e+03
385	5	20.8	374 35	W49075	Recombinant human Met	2.49e+02	458	4	16.7	37	16	R13952	CD16 glycosyl-phospha	2.22e+03
386	5	20.8	379 37	W83962	Recombinant human met	2.49e+02	459	4	16.7	38	3	R13952	A.niger Polygalacturo	2.22e+03
387	5	20.8	379 35	W49074	Recombinant human Met	2.49e+02	460	4	16.7	39	11	R60397	N-terminal fragment o	2.22e+03
388	5	20.8	379 35	W49073	Recombinant human Met	2.49e+02	461	4	16.7	49	27	W27692	Heavy chain site HCN3	2.22e+03
389	5	20.8	396 4	R22220	Sequence of secretory	2.49e+02	462	4	16.7	52	38	W88780	Polypeptide fragment	2.22e+03
390	5	20.8	396 4	R22216	Sequence of human int	2.49e+02	463	4	16.7	66	10	R61107	Bacillus firmus sodiu	2.22e+03
391	5	20.8	396 23	W10535	Leptin 1-167/IgG4 hin	2.49e+02	464	4	16.7	67	39	W89453	Hepatitis G virus PNF	2.22e+03
392	5	20.8	396 23	W10534	Leptin 1-167/IgG4 hin	2.49e+02	465	4	16.7	67	18	R90773	Hepatitis G virus PNF	2.22e+03
393	5	20.8	397 23	W22722	Human obesity protein	2.49e+02	466	4	16.7	69	3	R13348	P195 HCV antigen (195	2.22e+03
394	5	20.8	397 26	W24060	Human obesity protein	2.49e+02	467	4	16.7	79	39	W89454	Hepatitis G virus PNF	2.22e+03
395	5	20.8	399 19	W04218	Human urinary bladder	2.49e+02	468	4	16.7	79	18	R90798	Antigen encoded by PN	2.22e+03
396	5	20.8	399 23	W10536	Leptin 1-167/IgG1 hin	2.49e+02	469	4	16.7	84	38	W89863	Antigen 2 from clust	2.22e+03
397	5	20.8	401 23	W10537	Leptin 1-167/IgG1 hin	2.49e+02	470	4	16.7	90	39	W89613	Endometrial steroid b	2.22e+03
398	5	20.8	404 7	R34581	Mammalian serotonin r	2.49e+02	471	4	16.7	90	26	W35803	Human endometrial spe	2.22e+03
399	5	20.8	404 10	R53767	Yeast Mei 1 protein.	2.49e+02	472	4	16.7	91	5	R25405	ED-B.	2.22e+03
400	5	20.8	415 34	W69228	Human lysosomal stali	2.49e+02	473	4	16.7	96	6	R29962	TCR V-beta chains uti	2.22e+03
401	5	20.8	420 13	R75754	BAV3 E1B 56K protein.	2.49e+02	474	4	16.7	99	24	W27174	Human recombinant tru	2.22e+03
402	5	20.8	420 4	R22219	Sequence of secretory	2.49e+02	475	4	16.7	103	38	W89999	Expressed antigen for	2.22e+03
403	5	20.8	420 37	W82842	Human interleukin-5 r	2.49e+02	476	4	16.7	110	39	W89473	Hepatitis G virus epi	2.22e+03
404	5	20.8	420 4	R22215	Sequence of human int	2.49e+02	477	4	16.7	110	12	R66335	Human immunoglobulin	2.22e+03
405	5	20.8	421 5	R25064	Human IL-5 receptor a	2.49e+02	478	4	16.7	111	38	W82495	Ehrlichia sp. E80.4 p	2.22e+03
406	5	20.8	438 36	W49859	Thermotoga OCl/4V gly	2.49e+02	479	4	16.7	111	38	R24492	Ehrlichia sp. E46.4 p	2.22e+03
407	5	20.8	438 26	W34555	OC1/4 glycosidase 33G	2.49e+02	480	4	16.7	114	38	W90000	Expressed antigen for	2.22e+03
408	5	20.8	441 22	W20402	H. pylori cytoplasmic	2.49e+02	481	4	16.7	120	38	W70800	Amino acid sequence o	2.22e+03
409	5	20.8	442 22	W20624	H. pylori cytoplasmic	2.49e+02	482	4	16.7	121	21	W13823	Yeast transcription r	2.22e+03
410	5	20.8	445 10	R54782	Human brain serotonin	2.49e+02	483	4	16.7	123	38	W89930	Antigen 3 from cluste	2.22e+03
411	5	20.8	448 11	R57200	Rat 5HT6 receptor.	2.49e+02	484	4	16.7	123	17	R96097	Hepatitis E virus (Bu	2.22e+03
412	5	20.8	490 16	R90299	Yeast proteox-3 sequen	2.49e+02	485	4	16.7	123	8	R39305	Burma strain HEV ORF3	2.22e+03
413	5	20.8	490 31	W51257	Arabidopsis proto-por	2.49e+02	486	4	16.7	126	12	R63161	Mouse growth differen	2.22e+03
414	5	20.8	495 36	W80567	Thermostable glucose-	2.49e+02	487	4	16.7	129	22	W20471	H. pylori cytoplasmic	2.22e+03
415	5	20.8	505 34	W69249	Human p56-l protein.	2.49e+02	488	4	16.7	130	14	R79096	Human Fas ligand C-te	2.22e+03
416	5	20.8	509 18	R96086	Yeast calcineurin sub	2.49e+02	489	4	16.7	131	33	W64385	A. thaliana Apcr prot	2.22e+03
417	5	20.8	522 34	R770094	Sorting nexin 1.	2.49e+02	490	4	16.7	133	33	W64386	S. tuberosum Kpcr pro	2.22e+03
418	5	20.8	547 23	W22707	N-terminal truncated	2.49e+02	491	4	16.7	136	38	R88593	Secreted protein enco	2.22e+03
419	5	20.8	553 23	W22706	Arginyl-tRNA syntheta	2.49e+02	492	4	16.7	136	14	R88307	N-terminally deleted	2.22e+03
420	5	20.8	598 8	R39555	Sequence encoded by p	2.49e+02	493	4	16.7	136	29	W55239	H. pylori ORF 04gpl12	2.22e+03
421	5	20.8	755 20	W03550	Partial ORF-3 protein	2.49e+02	494	4	16.7	137	14	R79100	Human Fas ligand (par	2.22e+03
422	5	20.8	833 27	W32114	Streptococcus pneumonia	2.49e+02	495	4	16.7	138	14	R88305	N-terminally deleted	2.22e+03
423	5	20.8	864 17	R93021	Human glucagon degrad	2.49e+02	496	4	16.7	138	14	R79068	Mouse Fas ligand (par	2.22e+03
424	5	20.8	873 23	W19918	Mouse Ksr-1 (kinase s	2.49e+02	497	4	16.7	141	14	R88303	N-terminally deleted	2.22e+03
425	5	20.8	875 23	W19919	Human Ksr-1 (kinase s	2.49e+02	498	4	16.7	144	38	W88730	Secreted protein enco	2.22e+03
426	5	20.8	925 14	R79148	Human insulin recepto	2.49e+02	499	4	16.7	144	20	W09015	Immunogenic type F bo	2.22e+03
427	5	20.8	993 1	R05935	Secreted GPIIb subuni	2.49e+02	500	4	16.7	146	23	W11274	Bovine leptin for tre	2.22e+03
428	5	20.8	1026 14	R70121	IL5-R-GFP 130 fusion	2.49e+02	501	4	16.7	146	23	W11273	Porcine leptin for tr	2.22e+03
429	5	20.8	1475 2	R08221	Recombinant alpha amy	2.49e+02	502	4	16.7	146	32	W57254	Amino acid sequence o	2.22e+03
430	5	20.8	1493 5	R27640	Human calcium channel	2.49e+02	503	4	16.7	146	23	W08597	Properly folded obes1	2.22e+03
431	5	20.8	1577 16	R91047	Alpha-D-glucosyltrans	2.49e+02	504	4	16.7	146	28	W45467	Anti-obesity protein,	2.22e+03
432	5	20.8	1844 25	W18302	Photorehabdus luminesc	2.49e+02	505	4	16.7	147	24	W27166	Human recombinant OB	2.22e+03
433	5	20.8	1844 31	W56558	Toxin TcbAii, encoded	2.49e+02	506	4	16.7	147	26	W22434	Mouse Met-OB protein.	2.22e+03
434	5	20.8	2251 14	R71009	Human neuronal calciu	2.49e+02	507	4	16.7	150	39	W89790	Staphylococcus aureus	2.22e+03
435	5	20.8	2270 13	R69604	Calcium channel alpha	2.49e+02	508	4	16.7	151	10	R55161	Fragment of retinoic	2.22e+03
436	5	20.8	2270 14	R71010	Human neuronal calciu	2.49e+02	509	4	16.7	154	39	W90081	C. cellulovorans CBD-	2.22e+03
437	5	20.8	2304 1	R05707	Acetyl-CoA-carboxylas	2.49e+02	510	4	16.7	155	23	W19751	C. elegans inhibitor	2.22e+03
438	5	20.8	2354 25	W17871	Photorehabdus luminesc	2.49e+02	511	4	16.7	156	39	W90080	C. cellulovorans CBD-	2.22e+03
439	5	20.8	2504 31	W56557	Toxin TcbA, encoded b	2.49e+02	512	4	16.7	158	10	R53248	Fasciola hepatica cat	2.22e+03
440	5	20.8	3084 35	W50891	Mouse laminin A chain	2.49e+02	513	4	16.7	159	20	W07505	Mouse ob polypeptide	2.22e+03
441	5	20.8	3722 2	R10145	Cephalosporin antibio	2.49e+02	514	4	16.7	162	39	W90077	C. cellulovorans CbpA	2.22e+03
442	4	16.7	6 30	W58778	Human breast cancer g	2.22e+03	515	4	16.7	162	5	R25124	Hepatitis C virus ENV	2.22e+03
443	4	16.7	9 24	W28851	HTIV-1a,c derived pep	2.22e+03	516	4	16.7	163	3	R13402	Parvo virus B19 PAN-2	2.22e+03
444	4	16.7	9 18	R77927	Antigenic tbpl peptid	2.22e+03	517	4	16.7	165	22	W20621	H. pylori cytoplasmic	2.22e+03
445	4	16.7	9 30	W53080	Tbpl antigenic peptid	2.22e+03	518	4	16.7	167	21	W04315	Rat obese (ob) protei	2.22e+03
446	4	16.7	9 9	R47389	Pbpf-activity-display	2.22e+03	519	4	16.7	167	28	W41396	Chimeric anti-CEA ant	2.22e+03
447	4	16.7	11 23	W11809	Fas ligand antigen M5	2.22e+03	520	4	16.7	168	1	P91059	Nicotiana tabacum pat	2.22e+03
448	4	16.7	15 39	W67635	Multiple sclerosis se	2.22e+03	521	4	16.7	168	29	W42340	CTLA-4 extracellular	2.22e+03
449	4	16.7	15 1	R04093	Antigenic determinant	2.22e+03	522	4	16.7	169	14	R75509	N. alata arabinogalac	2.22e+03
450	4	16.7	20 24	W27008	Monomer subunit of er	2.22e+03	523	4	16.7	172	16	R82811	Rhipecephalus appendi	2.22e+03
451	4	16.7	22 29	W33963	CCR5 receptor antagon	2.22e+03	524	4	16.7	181	32	W56618	Human BY55 protein.	2.22e+03
452	4	16.7	22 29	W33996	CCR5 receptor antagon	2.22e+03	525	4	16.7	181	17	R89384	Barley ADP ribosylati	2.22e+03
453	4	16.7	23 10	R51684	Basophil granule prot	2.22e+03	526	4	16.7	182	39	W67960	Fragment of human sec	2.22e+03
454	4	16.7	28 5	R24868	Sequence of peptide f	2.22e+03	527	4	16.7	183	39	W84323	Gp4 protein of PRNSV	2.22e+03

528	4	16.7	184	26	W23069	Staphylothermus marin	2.22e+03
529	4	16.7	185	39	W82004	Human adult testis se	2.22e+03
530	4	16.7	186	1	R05742	3T6-R400 mutenin dihyd	2.22e+03
531	4	16.7	192	13	R69638	Hepatitis C virus env	2.22e+03
532	4	16.7	192	22	W24672	H. pylori secreted or	2.22e+03
533	4	16.7	192	13	R69637	Hepatitis C virus env	2.22e+03
534	4	16.7	192	13	R69639	Hepatitis C virus env	2.22e+03
535	4	16.7	193	7	R33985	HCT23 E1 protein.	2.22e+03
536	4	16.7	195	5	R26080	TAP.	2.22e+03
537	4	16.7	196	10	R50001	Platelet-derived Grow	2.22e+03
538	4	16.7	196	12	R68618	Human PDGF-A precursor	2.22e+03
539	4	16.7	196	3	R15644	PDGF 13-1 A-chain pre	2.22e+03
540	4	16.7	205	39	W73484	Grapevine leafroll vi	2.22e+03
541	4	16.7	208	39	W89431	Streptococcus pneumon	2.22e+03
542	4	16.7	209	11	R60211	Immunogenic fragment	2.22e+03
543	4	16.7	211	38	W67614	Hepatitis C virus E1	2.22e+03
544	4	16.7	211	15	R83558	Human modified 7B2 pr	2.22e+03
545	4	16.7	211	38	W87560	Human CTLA4 receptor	2.22e+03
546	4	16.7	211	7	R38918	Human Platelet Derive	2.22e+03
547	4	16.7	211	3	R15645	PDGF D-1 A-chain prec	2.22e+03
548	4	16.7	212	39	W82685	L. corniculatum type	2.22e+03
549	4	16.7	212	38	W81584	Human CTLA4 receptor	2.22e+03
550	4	16.7	212	39	W82678	X. cumberlandia type	2.22e+03
551	4	16.7	212	39	W82677	X. cumberlandia type	2.22e+03
552	4	16.7	215	39	W88238	HIV-1 co-receptor CCR	2.22e+03
553	4	16.7	215	16	R80783	Fibroblast growth fac	2.22e+03
554	4	16.7	217	17	R81945	Hepatitis C virus ant	2.22e+03
555	4	16.7	217	38	W80517	Insert of a HCV clone	2.22e+03
556	4	16.7	217	38	W80516	Hepatitis C virus ant	2.22e+03
557	4	16.7	217	38	W70829	Insert of a HCV clone	2.22e+03
558	4	16.7	217	38	W70828	Hepatitis C virus ant	2.22e+03
559	4	16.7	217	14	R75460	Mouse antibody H3-3 h	2.22e+03
560	4	16.7	218	39	W30669	TutB protein of tolu	2.22e+03
561	4	16.7	218	30	W47019	Arabidopsis thaliana	2.22e+03
562	4	16.7	219	31	W53597	Amino acid sequence o	2.22e+03
563	4	16.7	223	14	R75456	Mouse antibody FB3-2	2.22e+03
564	4	16.7	224	16	R88274	Papilloma virus major	2.22e+03
565	4	16.7	225	16	R88270	Papilloma virus major	2.22e+03
566	4	16.7	227	16	W13736	EHV-1 strain Ab4 gene	2.22e+03
567	4	16.7	228	21	R88275	Papilloma virus major	2.22e+03
568	4	16.7	231	21	W20162	H. pylori inner membr	2.22e+03
569	4	16.7	232	39	W89430	Streptococcus pneumon	2.22e+03
570	4	16.7	232	38	R89838	Protein encoded by cl	2.22e+03
571	4	16.7	232	39	W89890	Antigen 2 from cluste	2.22e+03
572	4	16.7	237	1	R06622	Hybrid (1,3-1,4)-pre-	2.22e+03
573	4	16.7	238	8	R38874	Sequence of haemagglu	2.22e+03
574	4	16.7	239	9	R45441	Mouse orphan receptor	2.22e+03
575	4	16.7	239	22	W14595	Aspergillus niger cel	2.22e+03
576	4	16.7	242	1	R04208	D protein- a recombin	2.22e+03
577	4	16.7	242	23	W22376	S. pneumoniae ORF27 p	2.22e+03
578	4	16.7	242	19	P95000	Bacillus subtilis lic	2.22e+03
579	4	16.7	243	22	W20501	H. pylori surface mem	2.22e+03
580	4	16.7	248	38	W89908	Antigen 1 from cluste	2.22e+03
581	4	16.7	248	34	W71031	Acid phosphatase enzy	2.22e+03
582	4	16.7	249	15	R77611	Humanised 5G1.1 VH +	2.22e+03
583	4	16.7	250	39	W84048	Rice OSMADS6 protein	2.22e+03
584	4	16.7	250	38	W89998	Expressed antigen for	2.22e+03
585	4	16.7	250	20	W06177	Murine A5B57 Fd fragm	2.22e+03
586	4	16.7	250	38	W89882	Antigen 1 from cluste	2.22e+03
587	4	16.7	251	28	W43920	Murine variable and h	2.22e+03
588	4	16.7	252	22	W20535	Helicobacter pylori c	2.22e+03
589	4	16.7	253	38	W81586	CTLA4/E7 fusion prote	2.22e+03
590	4	16.7	253	38	W87562	CTLA4-E7 fusion prote	2.22e+03
591	4	16.7	253	22	W15263	Cysteine protease Mch	2.22e+03
592	4	16.7	254	26	W29240	Human variant Fc-gamm	2.22e+03
593	4	16.7	255	28	W41394	Chimeric anti-CEA ant	2.22e+03
594	4	16.7	255	28	W41399	Humanised antibody 80	2.22e+03
595	4	16.7	260	28	W41409	Humanised heavy chain	2.22e+03
596	4	16.7	266	39	W90238	E. coli antibiotic ef	2.22e+03
597	4	16.7	268	8	R41691	AIGF #2.	2.22e+03
598	4	16.7	269	11	R60210	Immunogenic fragment	2.22e+03
599	4	16.7	271	38	W89833	Protein encoded by cl	2.22e+03
600	4	16.7	271	38	W89937	Antigen 1 from cluste	2.22e+03
601	4	16.7	271	5	R27727	HPV 18 E7 protein fra	2.22e+03
602	4	16.7	271	5	R27728	HPV 18 E6 protein fra	2.22e+03
603	4	16.7	273	39	W89748	Staphylococcus aureus	2.22e+03
604	4	16.7	281	18	R98104	Human Fas ligand deri	2.22e+03
605	4	16.7	282	19	W02703	G-protein coupled hum	2.22e+03
606	4	16.7	282	3	P60645	Mouse kidney cell Ban	2.22e+03
607	4	16.7	292	8	R44804	Human cyclin D3.	2.22e+03
608	4	16.7	292	16	R49418	Human glutamate recep	2.22e+03
609	4	16.7	292	30	W55719	H. pylori ORF 07e114	2.22e+03
610	4	16.7	293	39	W90237	E. coli antibiotic ef	2.22e+03
611	4	16.7	303	13	R71507	Dac gv (clone 259), m	2.22e+03
612	4	16.7	303	22	W15262	Apoptotic protease Mch	2.22e+03
613	4	16.7	306	39	W90236	E. coli antibiotic ef	2.22e+03
614	4	16.7	307	1	R04207	C13 protein - a recom	2.22e+03
615	4	16.7	307	11	R60208	Immunogenic fragment	2.22e+03
616	4	16.7	309	1	P94860	Expression plasmid pU	2.22e+03
617	4	16.7	315	16	R48713	G-protein coupled hum	2.22e+03
618	4	16.7	317	38	W81968	Mouse E124 protein.	2.22e+03
619	4	16.7	319	18	R98413	Mouse myeloma MOPC21	2.22e+03
620	4	16.7	321	39	W83375	Streptococcus pneumon	2.22e+03
621	4	16.7	321	18	R98412	Mouse immunoglobulin	2.22e+03
622	4	16.7	324	16	R92785	Soybean palmitoyl-ACP	2.22e+03
623	4	16.7	326	11	R57071	Fasciola hepatica cat	2.22e+03
624	4	16.7	327	32	W60223	Bacillus thuringiens	2.22e+03
625	4	16.7	327	22	W20901	H. pylori secreted or	2.22e+03
626	4	16.7	328	38	W81786	Tomato fructokinase (	2.22e+03
627	4	16.7	329	39	W86356	Partial human DNAX to	2.22e+03
628	4	16.7	329	38	W70801	Amino acid sequence o	2.22e+03
629	4	16.7	330	39	W89745	Staphylococcus aureus	2.22e+03
630	4	16.7	330	38	W85252	Maize PTE mature prot	2.22e+03
631	4	16.7	331	39	W67889	Human secreted protei	2.22e+03
632	4	16.7	332	38	W86253	Amino acid sequence o	2.22e+03
633	4	16.7	336	20	W03366	Toxoplasma gondii P30	2.22e+03
634	4	16.7	337	17	R91314	N. gonorrhoeae glycos	2.22e+03
635	4	16.7	337	32	W48734	Human RM3 seven trans	2.22e+03
636	4	16.7	337	30	W40100	Human herpesvirus 8 (	2.22e+03
637	4	16.7	338	38	W89928	Antigen 1 from cluste	2.22e+03
638	4	16.7	341	30	W53939	L.lactis MGL363 orfB	2.22e+03
639	4	16.7	341	31	W53944	L.lactis DBI341 orfB	2.22e+03
640	4	16.7	342	5	R25835	Human platelet activa	2.22e+03
641	4	16.7	344	39	W70838	Corn SMT enzyme.	2.22e+03
642	4	16.7	345	16	R94357	Human thymopoietin-ga	2.22e+03
643	4	16.7	351	8	R43685	Human kappa immunogl	2.22e+03
644	4	16.7	352	12	W23835	Human CC chemokine re	2.22e+03
645	4	16.7	352	12	R68812	Human monocyte PF4AR.	2.22e+03
646	4	16.7	352	39	W88232	HIV-1 co-receptor CCR	2.22e+03
647	4	16.7	352	24	W27125	Macaque chemokine rec	2.22e+03
648	4	16.7	352	15	R80757	Chemokine superfamily	2.22e+03
649	4	16.7	353	38	W89824	Protein encoded by cl	2.22e+03
650	4	16.7	355	32	W51745	Human C-C chemokine r	2.22e+03
651	4	16.7	355	18	W03377	CC-chemokine receptor	2.22e+03
652	4	16.7	357	34	W68134	Amino acid sequence o	2.22e+03
653	4	16.7	359	38	W85455	Secreted protein enco	2.22e+03
654	4	16.7	361	28	W34326	Tomato mottle virus A	2.22e+03
655	4	16.7	362	16	R48695	G-protein coupled hum	2.22e+03
656	4	16.7	363	38	W88563	Secreted protein enco	2.22e+03
657	4	16.7	363	22	W14139	Arabidopsis thaliana	2.22e+03
658	4	16.7	368	13	R68961	Xenorhabdus nematophi	2.22e+03
659	4	16.7	370	38	W88777	Polypeptide fragment	2.22e+03
660	4	16.7	378	27	W35864	Human FAS-ligand:IgG2	2.22e+03
661	4	16.7	380	29	W41502	Human cytokine/peptid	2.22e+03
662	4	16.7	384	39	W86318	Kidney injury associa	2.22e+03
663	4	16.7	385	32	W60043	Human MHC class I cha	2.22e+03
664	4	16.7	387	9	R49828	3-acylating enzyme.	2.22e+03
665	4	16.7	392	39	W73655	M. tuberculosis antig	2.22e+03
666	4	16.7	392	39	W73765	M. tuberculosis antig	2.22e+03
667	4	16.7	392	23	W18385	S. typhimurium prgh g	2.22e+03
668	4	16.7	393	12	R60700	Human ERK-5.	2.22e+03
669	4	16.7	394	2	R03754	Entire sequence of co	2.22e+03
670	4	16.7	394	3	P61708	[Glx358] alphas-antit	2.22e+03
671	4	16.7	394	3	P61713	[Phe358] alphas-antit	2.22e+03
672	4	16.7	394	3	P61712	[Ileu358] alphas-antit	2.22e+03
673	4	16.7	395	2	P83189	[Ala357, Arg358] Alph	2.22e+03

674	4	16.7	397	4	R1328	Sequence of protein N	2.22e+03	502	38	W81585	CTLA4/p97 fusion prot	2.22e+03
675	4	16.7	398	39	W87566	A cyclin related prot	2.22e+03	502	38	W87561	CTLA4-p97 fusion prot	2.22e+03
676	4	16.7	402	8	R1310	Transcriptase.	2.22e+03	507	36	W83124	PTIR50 Arginine spe	2.22e+03
677	4	16.7	402	2	R06495	Beta 3 adrenergic rec	2.22e+03	509	35	W49861	Thermococcus 9N2-31B/	2.22e+03
678	4	16.7	404	9	R48674	Putative peach polyga	2.22e+03	510	3	R15354	Protein deduced from	2.22e+03
679	4	16.7	405	2	P70165	Beta-isopropylmalate	2.22e+03	514	9	R47342	Mammalian synaptic ve	2.22e+03
680	4	16.7	406	10	R57077	Sialyltransferase.	2.22e+03	514	38	W81976	Ehrlichia sp. W20.2 p	2.22e+03
681	4	16.7	406	39	W26880	Streptomyces sp. type	2.22e+03	514	38	W73378	CSB2 protein sequence	2.22e+03
682	4	16.7	409	5	R25693	Heat resistant carbam	2.22e+03	526	32	W44845	Cyclomaltodextrin glu	2.22e+03
683	4	16.7	415	16	R92787	Canola palmitoyl-ACP-	2.22e+03	526	34	W69963	Aeromonas caviae xyla	2.22e+03
684	4	16.7	416	38	W88745	Secreted protein enco	2.22e+03	530	9	R46679	CD4-IG2 chimeric hea	2.22e+03
685	4	16.7	418	3	P30877	Sequence encoded by h	2.22e+03	530	5	R26783	CD4-IG2 chimeric hea	2.22e+03
686	4	16.7	419	38	W83397	Maize DnaJ clone prot	2.22e+03	537	39	W87796	Tripeptidyl aminopept	2.22e+03
687	4	16.7	419	38	W86251	Maize PTE enzyme.	2.22e+03	539	29	W37128	Hepatitis C virus opt	2.22e+03
688	4	16.7	421	24	W00936	Adenovirus 1 subtype	2.22e+03	540	39	W89794	Staphylococcus aureus	2.22e+03
689	4	16.7	421	17	R97594	Adenovirus 1 subtype	2.22e+03	540	39	W89777	Staphylococcus aureus	2.22e+03
690	4	16.7	427	1	P80667	Mutant aroA gene enco	2.22e+03	541	39	W67820	Human secreted protei	2.22e+03
691	4	16.7	427	5	R77794	EPSP synthase wild ty	2.22e+03	543	29	W41779	dsR gene product.	2.22e+03
692	4	16.7	431	20	W9014	Immunogenic type F bo	2.22e+03	547	28	W43013	Truncated transferrin	2.22e+03
693	4	16.7	435	12	R58689	Rat REC20 serotonin r	2.22e+03	547	26	W22053	Choline oxidase deriv	2.22e+03
694	4	16.7	442	21	W10550	IgG1 polypeptide.	2.22e+03	548	39	W89189	Alternatively spliced	2.22e+03
695	4	16.7	442	21	W10551	IgG1 polypeptide with	2.22e+03	549	38	W70898	Acetyl-coenzyme A tra	2.22e+03
696	4	16.7	443	21	W13564	Humanised anti-L-sele	2.22e+03	551	39	W73480	Grapevine leafroll vi	2.22e+03
697	4	16.7	443	8	R43895	hEL-nl.	2.22e+03	552	6	R30775	pH52-8.0 humanised mu	2.22e+03
698	4	16.7	446	39	W83374	Streptococcus pneumon	2.22e+03	557	39	W81936	A. thaliana trehalase	2.22e+03
699	4	16.7	446	32	W37905	Mycobacterium tubercu	2.22e+03	558	38	W73379	CSB2 protein sequence	2.22e+03
700	4	16.7	447	22	W01622	Oyster pearl protein.	2.22e+03	560	38	W79954	Truncated adenovirus	2.22e+03
701	4	16.7	448	17	R5315	Human hepatoma retino	2.22e+03	562	39	W85602	Hexosaminidase enzyme	2.22e+03
702	4	16.7	449	32	W86739	Orpinomyces cellulase	2.22e+03	563	10	R53923	Acyl CoA synthetase.	2.22e+03
703	4	16.7	453	12	R3790	Aspergillus aculeatus	2.22e+03	565	1	R04943	Equine hemagglutinin	2.22e+03
704	4	16.7	454	17	R93187	Thymopoleitin-beta.	2.22e+03	566	34	W68405	SIV strain HIN1 haema	2.22e+03
705	4	16.7	459	8	R42066	Human anti-HBS heavy	2.22e+03	568	38	W79953	Truncated adenovirus	2.22e+03
706	4	16.7	461	16	R92220	Calcitonin gene-relat	2.22e+03	570	1	P70710	Equine influenza viru	2.22e+03
707	4	16.7	461	16	R75365	Phytase.	2.22e+03	570	1	P94046	l-plastin protein der	2.22e+03
708	4	16.7	462	38	W81964	Human PAR-alpha prote	2.22e+03	573	38	W79955	Truncated adenovirus	2.22e+03
709	4	16.7	462	2	R10547	Murine Retinoic Acid	2.22e+03	573	38	W79952	Truncated adenovirus	2.22e+03
710	4	16.7	462	21	W44930	2A2 (Chimeric) human	2.22e+03	575	9	R47175	Sequence of polyepit	2.22e+03
711	4	16.7	463	21	W49434	3F4 (Chimeric) human	2.22e+03	575	6	R30045	gE gene deletion from	2.22e+03
712	4	16.7	464	39	W83041	Anti-Fas Mab HFE7A he	2.22e+03	581	38	W75419	Mutant adenovirus typ	2.22e+03
713	4	16.7	465	26	W37438	Rat islet glucokinase	2.22e+03	581	38	W79950	Mutant adenovirus typ	2.22e+03
714	4	16.7	465	8	R50195	Human wild type gluco	2.22e+03	581	38	W79949	Mutant adenovirus typ	2.22e+03
715	4	16.7	466	22	W7820	Pseudomonas putida am	2.22e+03	581	38	W79905	Adenovirus type fibre	2.22e+03
716	4	16.7	466	4	A40032	Gamma anti-carcinomb	2.22e+03	581	38	W82732	Adenovirus wild-type	2.22e+03
717	4	16.7	467	38	W82609	Murine Smad2 S465A mu	2.22e+03	581	38	W79951	Mutant adenovirus typ	2.22e+03
718	4	16.7	467	38	W82607	Murine Smad2 protein.	2.22e+03	585	38	W73315	Parathyroid hormone r	2.22e+03
719	4	16.7	467	39	W82595	Mouse Smad2 protein.	2.22e+03	586	2	R07274	Reverse transcriptase	2.22e+03
720	4	16.7	467	38	W73400	Human secreted protei	2.22e+03	586	13	R73935	msDNA-Ec67 reverse-tr	2.22e+03
721	4	16.7	467	38	W82612	Murine Smad2 S464A/S4	2.22e+03	587	8	R39553	Deduced amino acid se	2.22e+03
722	4	16.7	467	38	W82611	Murine Smad2 S464A/S4	2.22e+03	595	39	W86003	Anti-5T4 single chain	2.22e+03
723	4	16.7	467	38	W82614	Murine Smad2 S464A/S4	2.22e+03	597	19	W00334	Bacillus thuringiens	2.22e+03
724	4	16.7	467	38	W82613	Murine Smad2 S465A/S4	2.22e+03	599	39	W81974	Lactococcus lactis Lt	2.22e+03
725	4	16.7	467	38	W82610	Murine Smad2 S467A mu	2.22e+03	605	18	R99673	Receptor for hyaluron	2.22e+03
726	4	16.7	467	38	W82608	Murine Smad2 S464A mu	2.22e+03	617	15	R79036	Infectious bovine rhi	2.22e+03
727	4	16.7	467	28	W29459	Mad-related human JVI	2.22e+03	619	5	R27651	Human calcium channel	2.22e+03
728	4	16.7	467	23	W14927	Human gamma-4pE heavy	2.22e+03	620	23	W22225	S. cerevisiae origin	2.22e+03
729	4	16.7	470	39	W83036	Anti-Fas humanised an	2.22e+03	621	29	W55645	H. pylori ORF 06p106	2.22e+03
730	4	16.7	470	39	W83037	Anti-Fas humanised an	2.22e+03	621	28	W02284	Candida albicans leuk	2.22e+03
731	4	16.7	470	34	W68135	Amino acid sequence o	2.22e+03	623	18	W03140	Human IgG CHI-hinge-C	2.22e+03
732	4	16.7	470	9	R47118	Mtr protein of Neuros	2.22e+03	627	8	R41229	GAT-3 GABA transporte	2.22e+03
733	4	16.7	472	18	R93166	Anti-rhesus D recombi	2.22e+03	627	9	R49590	GAT-B transporter.	2.22e+03
734	4	16.7	475	21	W16421	Human anti-RSV monocl	2.22e+03	627	2	R11491	Leucocyte plastin.	2.22e+03
735	4	16.7	476	21	W01822	Primates anti-human	2.22e+03	631	18	R99675	RHAMM 1-2a isoform.	2.22e+03
736	4	16.7	476	32	W63765	Macaque primatized 16	2.22e+03	638	38	W86014	R. rubrum poly-beta-h	2.22e+03
737	4	16.7	478	32	W63763	Macaque primatized 7B	2.22e+03	639	2	R11271	B. stearothermophilus	2.22e+03
738	4	16.7	483	25	W93951	Filamin-like beta 7 i	2.22e+03	644	39	W82318	Human 7-transmembrane	2.22e+03
739	4	16.7	484	13	R75380	Natural resistance-as	2.22e+03	644	26	W34814	Novel CryIIIA mutant	2.22e+03
740	4	16.7	485	2	R1919	Reverse transcriptase	2.22e+03	644	26	W34811	Sequence of toxin whi	2.22e+03
741	4	16.7	493	39	W84068	Human mitofusin Fzo h	2.22e+03	644	21	R56697	Novel CryIIIA mutant	2.22e+03
742	4	16.7	494	30	W47017	Arabidopsis thaliana	2.22e+03	644	11	R56697	CryIIIA insecticidal	2.22e+03
743	4	16.7	497	2	P70572	Glucanase gene pro	2.22e+03	647	28	W43014	Truncated transferrin	2.22e+03
744	4	16.7	499	12	R60609	Tobamovirus replicati	2.22e+03	660	17	R96970	Bacterial transferrin	2.22e+03
745	4	16.7	499	5	R27574	The ABF-B from A. nig	2.22e+03	661	12	R63226	Mycoplasma gallisepti	2.22e+03
746	4	16.7	501	20	W06826	Marek's disease virus	2.22e+03	667	26	W22009	Glucan elicitor recep	2.22e+03

820	4	16.7	671	31	W53458	Fly latheo protein se	2.22e+03	893	4	16.7	944	2	R06427	Protein encoded by se	2.22e+03
821	4	16.7	671	3	R14548	T cell transmembrane	2.22e+03	894	4	16.7	951	38	W70798	Human gpi30-C-gamma-1	2.22e+03
822	4	16.7	682	21	W17649	CGase variant 87-94	2.22e+03	895	4	16.7	958	1	P81181	Sequence of glucosyl	2.22e+03
823	4	16.7	683	21	W17625	Thermoanaerobacter CG	2.22e+03	896	4	16.7	970	15	R85089	EPH-like receptor pro	2.22e+03
824	4	16.7	683	21	W17636	Thermoanaerobacter CG	2.22e+03	897	4	16.7	979	39	W83068	TutC protein of tolu	2.22e+03
825	4	16.7	683	21	W17641	Thermoanaerobacter CG	2.22e+03	898	4	16.7	984	5	R27529	Plasmodium falciparum	2.22e+03
826	4	16.7	683	21	W17630	Thermoanaerobacter CG	2.22e+03	899	4	16.7	1018	22	W06485	Rat contactin ligand	2.22e+03
827	4	16.7	683	21	W17624	Thermoanaerobacter CG	2.22e+03	900	4	16.7	1019	8	R38096	Pyrococcus sp DNA pol	2.22e+03
828	4	16.7	683	21	W17623	Thermoanaerobacter CG	2.22e+03	901	4	16.7	1027	39	W89806	Staphylococcus aureus	2.22e+03
829	4	16.7	683	21	W17622	Thermoanaerobacter CG	2.22e+03	902	4	16.7	1036	30	W53280	Human phospholipase D	2.22e+03
830	4	16.7	684	21	W17580	Thermoanaerobacter CG	2.22e+03	903	4	16.7	1039	38	W85014	Stats-green flouresce	2.22e+03
831	4	16.7	684	21	W17588	Thermoanaerobacter CG	2.22e+03	904	4	16.7	1045	38	W85050	Stats-green flouresce	2.22e+03
832	4	16.7	684	21	W17589	Thermoanaerobacter CG	2.22e+03	905	4	16.7	1053	15	R88578	Chicken focal adhesio	2.22e+03
833	4	16.7	685	5	R25911	Macrophage tissue tra	2.22e+03	906	4	16.7	1061	24	W26370	Human steroid recepto	2.22e+03
834	4	16.7	686	3	P61518	Sequence of cyclomalt	2.22e+03	907	4	16.7	1075	34	W64472	Human secreted protei	2.22e+03
835	4	16.7	691	2	R04711	Sequence of guinea pi	2.22e+03	908	4	16.7	1079	39	W89566	Rat parathyroid calci	2.22e+03
836	4	16.7	698	39	W83493	4D5 Fab molecule expr	2.22e+03	909	4	16.7	1085	39	W89563	Bovine parathyroid ca	2.22e+03
837	4	16.7	707	39	W84052	Human V3 loop HIV rec	2.22e+03	910	4	16.7	1085	28	W38272	B lymphocyte membrane	2.22e+03
838	4	16.7	708	38	W83394	Human protein-coupled	2.22e+03	911	4	16.7	1087	2	R11139	Truncated tie recepto	2.22e+03
839	4	16.7	714	27	W35316	M. catarrhalis RI tra	2.22e+03	912	4	16.7	1094	8	R39821	Human G-protein regul	2.22e+03
840	4	16.7	718	38	W85024	Erk2-green flourescen	2.22e+03	913	4	16.7	1101	39	W90089	Rat orphan tyrosine k	2.22e+03
841	4	16.7	719	38	W85012	Smad2-green flouresce	2.22e+03	914	4	16.7	1101	38	W83160	Porcine G-protein reg	2.22e+03
842	4	16.7	719	38	W85012	Modified protective a	2.22e+03	915	4	16.7	1102	39	W90083	Neurofibromatosis typ	2.22e+03
843	4	16.7	733	7	R38278	NANB hepatitis virus	2.22e+03	916	4	16.7	1103	8	R39631	CIITA (class II trans	2.22e+03
844	4	16.7	738	5	R24396	Prod. Of the S gene o	2.22e+03	917	4	16.7	1106	37	W80312	TIIM protein.	2.22e+03
845	4	16.7	749	23	W22479	Plasmodium ebl-1.	2.22e+03	918	4	16.7	1122	21	W14052	Human orphan receptor	2.22e+03
846	4	16.7	750	2	R10458	FB29-encoded chicken	2.22e+03	919	4	16.7	1124	9	R45440	CIITA (class II trans	2.22e+03
847	4	16.7	750	3	R14048	C-ski protein.	2.22e+03	920	4	16.7	1130	37	W80313	GST-SEP fusion protei	2.22e+03
848	4	16.7	763	24	W26407	Matrix/scaffold-assoc	2.22e+03	921	4	16.7	1140	16	R81731	Bacillus thuringiens	2.22e+03
849	4	16.7	764	38	W81784	P. putida cis/trans i	2.22e+03	922	4	16.7	1167	20	W10653	T.thermophilus nitrat	2.22e+03
850	4	16.7	767	10	R55709	Glycosyltransferase h	2.22e+03	923	4	16.7	1195	39	W75420	Urease protein.	2.22e+03
851	4	16.7	793	32	W37911	Rattus norvegicus Smo	2.22e+03	924	4	16.7	1207	37	W80311	TIIP20.	2.22e+03
852	4	16.7	794	15	R88199	Sheep mammary gland f	2.22e+03	925	4	16.7	1216	21	W14497	86Q3a toxin.	2.22e+03
853	4	16.7	797	38	W81963	Human myl/RAR-alpha f	2.22e+03	926	4	16.7	1230	21	W15764	MEK5 protein.	2.22e+03
854	4	16.7	813	39	W84307	A human CDC28-#46 RNA	2.22e+03	927	4	16.7	1245	21	W13872	Adenyl cyclase-1 (acy	2.22e+03
855	4	16.7	830	37	R80314	CIITA (class II trans	2.22e+03	928	4	16.7	1247	39	W73535	DNA polymerase 385 kd	2.22e+03
856	4	16.7	839	17	R81825	Human trkC receptor p	2.22e+03	929	4	16.7	1253	39	W88524	Bacillus subtilis srf	2.22e+03
857	4	16.7	843	38	W85030	Green fluorescent pro	2.22e+03	930	4	16.7	1263	26	W22052	H. contortus pop-A pr	2.22e+03
858	4	16.7	845	18	R77890	Bacterial transferrin	2.22e+03	931	4	16.7	1274	7	R34714	Bacillus thuringiens	2.22e+03
859	4	16.7	851	18	W03166	Human STAT2.	2.22e+03	932	4	16.7	1275	39	W82594	Human angiotensin con	2.22e+03
860	4	16.7	852	2	R04029	Env gene product of c	2.22e+03	933	4	16.7	1275	39	W82594	Aradiposis thaliana	2.22e+03
861	4	16.7	853	38	R80531	Protein tyrosine phos	2.22e+03	934	4	16.7	1275	39	W82594	Bovine Coronavirus E2	2.22e+03
862	4	16.7	854	36	W72113	HSV-2 strain SRS Cont	2.22e+03	935	4	16.7	1289	12	R58632	Procoagulant-active h	2.22e+03
863	4	16.7	856	10	R12449	Fiv PET-F14 envelope	2.22e+03	936	4	16.7	1306	1	R04111	Mouse WRN helicase pr	2.22e+03
864	4	16.7	856	23	W08436	Feline immunodeficien	2.22e+03	937	4	16.7	1321	17	R84823	Neuronal apoptosis in	2.22e+03
865	4	16.7	858	28	W23314	Human prostatic protei	2.22e+03	938	4	16.7	1363	2	R11061	E2 protein of Canine	2.22e+03
866	4	16.7	860	8	R42962	Bovine MTP.	2.22e+03	939	4	16.7	1383	28	W33228	CCV-C54 spike protein	2.22e+03
867	4	16.7	870	10	R51246	Composite sequence of	2.22e+03	940	4	16.7	1383	28	W33227	FECV/FIPV chimeric sp	2.22e+03
868	4	16.7	870	23	W19725	Truncated cell cycle	2.22e+03	941	4	16.7	1401	32	W59454	Prod. of the S gene o	2.22e+03
869	4	16.7	875	13	R73053	Peptidyl C-terminal a	2.22e+03	942	4	16.7	1403	23	W20032	Prod. of the S gene o	2.22e+03
870	4	16.7	875	13	P94854	C-terminal prepro-C-t	2.22e+03	943	4	16.7	1453	6	R31038	Human N-methyl-D-aspa	2.22e+03
871	4	16.7	878	38	W81973	Corn pullulanase prot	2.22e+03	944	4	16.7	1453	5	R27820	p190 protein.	2.22e+03
872	4	16.7	885	23	W22230	K. lactis origin of r	2.22e+03	945	4	16.7	1454	9	R42472	Borna disease virus p	2.22e+03
873	4	16.7	890	39	W73534	MEK4 protein.	2.22e+03	946	4	16.7	1454	5	R24395	Enzyme involved in ei	2.22e+03
874	4	16.7	894	8	R50007	Human MTP.	2.22e+03	947	4	16.7	1454	5	R24397	Pre-pro-cobra C3 prot	2.22e+03
875	4	16.7	901	2	R11994	Glutamate receptor 6.	2.22e+03	948	4	16.7	1474	38	W85136	Bacterial amylase A-1	2.22e+03
876	4	16.7	902	9	R48951	AMPA-binding human G1	2.22e+03	949	4	16.7	1474	3	R43253	Sequence of Heliothis	2.22e+03
877	4	16.7	903	11	R60183	PA(1-725)----Human C	2.22e+03	950	4	16.7	1513	8	R43253	Borna disease virus p	2.22e+03
878	4	16.7	903	16	R87007	Hyperthermostable pro	2.22e+03	951	4	16.7	1608	19	R98619	Enzyme involved in ei	2.22e+03
879	4	16.7	904	8	R41779	Glycoprotein B (gB2).	2.22e+03	952	4	16.7	1634	9	R42452	Bacterial amylase A-1	2.22e+03
880	4	16.7	904	26	R44553	Herpes simplex virus	2.22e+03	953	4	16.7	1651	12	R63222	Sequence of Heliothis	2.22e+03
881	4	16.7	904	8	R41778	Glycoprotein B (gB1).	2.22e+03	954	4	16.7	1684	3	R14948	Borna disease virus p	2.22e+03
882	4	16.7	905	13	R14565	HSVgB polypeptide.	2.22e+03	955	4	16.7	1704	9	R49657	Rat Munc13-1.	2.22e+03
883	4	16.7	905	11	R63068	Human EAA3b excitator	2.22e+03	956	4	16.7	1711	19	R98605	Chicken cytotactin.	2.22e+03
884	4	16.7	908	9	R48220	N.meningitidis IM2394	2.22e+03	957	4	16.7	1763	39	W83431	Urea amidolyase.	2.22e+03
885	4	16.7	908	9	R45367	Excitatory amino acid	2.22e+03	958	4	16.7	1810	16	R45863	Mutant protease (A137	2.22e+03
886	4	16.7	911	2	R10333	Deduced sequence of t	2.22e+03	959	4	16.7	1835	8	R42839	Mutant protease (K138	2.22e+03
887	4	16.7	912	18	R77888	Bacterial transferrin	2.22e+03	960	4	16.7	1962	2	R10557	Human neuronal calciu	2.22e+03
888	4	16.7	912	30	W53044	H. influenzae strain	2.22e+03	961	4	16.7	1962	2	R10560	Human Factor VIII.c	2.22e+03
889	4	16.7	913	30	W53042	H. influenzae strain	2.22e+03	962	4	16.7	2161	14	P71001	Human Factor VIII.	2.22e+03
890	4	16.7	914	30	W53048	H. influenzae strain	2.22e+03	963	4	16.7	2332	2	P71728	Procoagulant-active h	2.22e+03
891	4	16.7	914	3	R15785	B.thuringiensis toxin	2.22e+03	964	4	16.7	2332	8	R43257		
892	4	16.7	930	39	W75426	Human aggrscan degrad	2.22e+03	965	4	16.7	2332	28	W33224		

```

966 4 16.7 2344 24 W11367 Active Factor VIII:C 2.22e+03
967 4 16.7 2345 24 W11403 Active Factor VIII:C 2.22e+03
968 4 16.7 2347 24 W11466 Active Factor VIII:C 2.22e+03
969 4 16.7 2347 24 W11467 Active Factor VIII:C 2.22e+03
970 4 16.7 2347 24 W11402 Active Factor VIII:C 2.22e+03
971 4 16.7 2347 24 W11379 Active Factor VIII:C 2.22e+03
972 4 16.7 2349 24 W11401 Active Factor VIII:C 2.22e+03
973 4 16.7 2349 24 W11424 Active Factor VIII:C 2.22e+03
974 4 16.7 2349 24 W11365 Active Factor VIII:C 2.22e+03
975 4 16.7 2349 24 W11420 Active Factor VIII:C 2.22e+03
976 4 16.7 2349 24 W11373 Active Factor VIII:C 2.22e+03
977 4 16.7 2349 24 W11465 Active Factor VIII:C 2.22e+03
978 4 16.7 2350 24 W11380 Active Factor VIII:C 2.22e+03
979 4 16.7 2351 24 W10590 Factor VIII:C (Phe270 2.22e+03
980 4 16.7 2351 24 W10592 Factor VIII:C (Tyr178 2.22e+03
981 4 16.7 2351 24 W10591 Factor VIII:C (Phe552 2.22e+03
982 4 16.7 2351 18 W00465 Factor-VIII. 2.22e+03
983 4 16.7 2352 24 W11400 Active Factor VIII:C 2.22e+03
984 4 16.7 2352 24 W11388 Active Factor VIII:C 2.22e+03
985 4 16.7 2352 24 W11464 Active Factor VIII:C 2.22e+03
986 4 16.7 2352 24 W11463 Active Factor VIII:C 2.22e+03
987 4 16.7 2352 24 W11374 Active Factor VIII:C 2.22e+03
988 4 16.7 2391 10 R53694 Carbamoyl-phosphate-s 2.22e+03
989 4 16.7 2446 11 R60021 Fibrinogen-alpha. 2.22e+03
990 4 16.7 2458 2 R04031 Full length T4 encode 2.22e+03
991 4 16.7 2485 11 R59921 RAS associated GAP NF 2.22e+03
992 4 16.7 2549 16 R81730 Sirolimus effector pr 2.22e+03
993 4 16.7 2639 39 W74476 Grapevine leafroll vi 2.22e+03
994 4 16.7 2816 7 R34009 HCV-1 polyprotein. 2.22e+03
995 4 16.7 2818 21 W13280 Human neurofibromin. 2.22e+03
996 4 16.7 2989 22 W19688 ATM mutant 1407del201 2.22e+03
997 4 16.7 3011 7 R34468 Encoded by full-lengt 2.22e+03
998 4 16.7 3011 6 R31621 Hepatitis C virus (HC 2.22e+03
999 4 16.7 3056 22 W19699 ATM mutant Leu43Pro. 2.22e+03
1000 4 16.7 4544 9 R47861 Alpha 2-Macroglobulin 2.22e+03

```

## ALIGNMENTS

```

RESULT 1
ID W52839 standard; Protein; 109 AA.
AC W52839;
DE 10-JUL-1998 (first entry)
KW Secreted protein encoded by clone AY160.2.
KW Clone AY160.2; secreted protein; chromosome identification;
KW identification; genetic disorder; nutritional use; cell proliferation;
KW cell differentiation; haematopoiesis regulating activity; tissue growth;
KW anti-inflammatory; tumour inhibition.
OS Homo sapiens.
PN W09807855-A2.
PD 26-FEB-1998.
PF 20-AUG-1997; U14641.
PR 23-AUG-1996; US-701931.
PI (GENE) GENETICS INST INC.
PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA,
PI Spaulding V, Treacy M;
DR WPI: 98-169159/15.
DR N-PSDB: V21165.
PT Nucleic acid encoding secreted protein from human cells - useful
PT e.g. as immuno-modulators, anti-tumour agents, promoters of tissue
PT growth, haemostatic and thrombolytic agents etc.
PS Claim 20; Page 48; 63pp; English.
CC The present sequence represents secreted protein encoded by clone
CC AY160.2. This clone is a full length clone, isolated from a human adult
CC retina cDNA library, using methods which are selective for cDNAs
CC encoding secreted proteins. The products of clone AY160.2 may as
CC tissue/molecular weight markers, for chromosome identification,
CC to identify possible genetic disorders, and to isolate new related DNA.
CC They can also be used for nutritional uses, cytokine and cell
CC proliferation, cell differentiation activity, immune stimulating or
CC suppressing activity, haematopoiesis regulating activity, tissue growth
CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombolytic activity, receptor/ligand activity,

```

```

CC anti-inflammatory activity, tumour inhibition activity or other
CC activities.
SQ Sequence 109 AA;
Query Match 25.0%; Score 6; DB 30; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.33e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 81 vtglye 86
Qy 10 VTGLYE 15
|||||
RESULT 2
ID W75916 standard; Protein; 341 AA.
AC W75916;
DE Human p58 receptor protein.
KW p58 receptor protein; human; natural killer cell; NK; MHC;
KW cell surface molecule; major histocompatibility complex.
OS Homo sapiens.
PN US5808028-A.
PD 15-SEP-1998.
PR 24-MAY-1994; US-248628.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Long FO, Waetmann N;
PI WPI: 98-520214/44.
DR N-PSDB: V54656.
PT DNA encoding human p58 receptor protein - useful for producing
PT recombinant protein and for detecting nucleic acid encoding the p58
PT receptor and for screening for agonists and antagonists
PS Claim 2; Columns 11-16; 13pp; English.
CC This represents a human p58 receptor protein. A host cell can be
CC transformed or transfected with a vector containing the p58 receptor
CC protein encoding nucleic acid for the recombinant production of the
CC protein. The p58 receptor of natural killer (NK) cells is a family of
CC cell surface molecules and is involved in the recognition of target cells
CC by NK cells. The p58 receptor contributes to the specificity of NK cells
CC in their recognition of major histocompatibility complex (MHC) class I
CC molecules on target cells. The present invention provides methods for
CC detecting nucleic acid encoding p58 receptor protein as well as the
CC protein in biological samples. Methods for screening for drugs capable
CC of acting as agonists or antagonists to the p58 receptor protein are
CC also provided.
SQ Sequence 341 AA;
Query Match 25.0%; Score 6; DB 35; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.33e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 123 tglyek 128
Qy 11 TGLYER 16
|||||
RESULT 3
ID R76097 standard; peptide; 17 AA.
AC R76097;
DE 02-FEB-1996 (first entry)
DE Omega conotoxin GVIA-derived peptide 180.
KW Omega conotoxin; marine snail; Conus; voltage-gated Ca channel blocker;
KW synaptosome; membrane; fish electric organ; mammalian brain; ischaemia;
KW binding protein; binding affinity; stroke.
OS Synthetic.
FT Key Location/Qualifiers
FT disulfide_bond 1..16
FT modified_site 4
FT /label= 4-Hyp
FT modified_site 11
FT /label= 4-Hyp
FT modified_site 17
FT /note= "amidated C-terminus"

```

PN US5424218-A.  
PD 13-JUN-1995.  
PF 22-NOV-1989; 440094.  
PR 02-NOV-1989; US-440094.  
PR 02-AUG-1990; US-561766.  
PR 23-MAR-1992; US-855269.  
PR 04-NOV-1993; US-147714.  
PA (NEUR-) NEUREX CORP.  
PI Bither RS, Bowersox SS, Fox JA, Miljanich GP, Valentino KL;  
PI Yamashiro DH;  
DR WPI: 95-223694/29.  
PT Identifying cpds. able to reduce neuronal damage caused by ischemia  
PT - by measuring their affinity to omega conotoxin MVIIA binding site  
PT and ability e.g. to inhibit voltage gated calcium channels  
PS Disclosure: Fig 3: 31pp; English.  
CC The peptides R76096-109 are synthetic peptides derived from the sequences  
CC of the naturally occurring omega conotoxin (OCT) peptides GVIA (R76091) or  
CC MVIIA (R76089) from marine snails of the Conus genus. The OCT peptides  
CC act as voltage-gated Ca channel blockers by binding to a 210 kD protein  
CC from synaptosomal membrane preparations from fish electric organ or  
CC mammalian brains. The peptides and their synthesised fragments can be  
CC used to screen for compounds that bind to the OCT binding protein, by  
CC displacing a high affinity labelled OCT, such as MVIIA, from a  
CC synaptosomal membrane preparation. The compounds should have binding  
CC affinities and activities at least equal to those of the natural peptides  
CC (Ki 0.44-324 nM). The screened compounds are potentially useful in  
CC treating ischemic conditions, esp. stroke, and can reduce sec.  
CC anatomical and functional damage associated with those conditions.  
SQ Sequence 17 AA;

Query Match 20.8%; Score 5; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 ytkrc 16  
|||||  
QY 1 YTKRC 5

RESULT 4  
ID R32785 standard; peptide: 17 AA.  
AC R32785;  
DT 28-JUN-1993 (first entry)  
DE GVIA omega conotoxin derived peptide fragment 180.  
KW OCT: neuronal damage reduction; ischemia; secondary damage; stroke.  
OS Synthetic.  
PN US5189020-A.  
PD 23-FEB-1993.  
PF 02-AUG-1990; 561766.  
PR 22-NOV-1989; US-440094.  
PR 02-AUG-1990; US-561766.  
PA (NEUR-) NEUREX CORP.  
PI Bither RS, Bowersox SS, Fox JA, Miljanich GP, Tsubokawa M,  
PI Valentino KL, Yamashiro DH;  
DR WPI: 93-085564/10.  
PT Reducing neuronal damage due to ischemia - involves using omega  
PT conotoxin peptide or fragment  
PS Disclosure: Fig 3: 32pp; English.  
CC The sequence is that of GVIA omega conotoxin (OCT) derived peptide  
CC fragment 180 which can bind to an OCT binding protein, inhibit voltage  
CC -gated calcium currents selectively in neuronal tissue and inhibit  
CC neuronal transmitter release selectively in neuronal tissue. These  
CC properties all occur within the range of those of MVIIB, GVIA, RVIA,  
CC or pref. MVIIA and GVIA OCTs. The peptide can be used in reducing or  
CC preventing both anatomical and functional secondary damage related  
CC to ischemia, generally as associated with stroke.  
SQ Sequence 17 AA;

Query Match 20.8%; Score 5; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 ytkrc 16

QY 1 YTKRC 5  
|||||  
RESULT 5  
ID W95566 standard; protein: 27 AA.  
AC W95566;  
DT 29-MAR-1999 (first entry)  
DE Omega-conopeptide GVIA/SNX-124.  
KW Omega-conopeptide; peptide toxin; snail; calcium channel blocker;  
KW analgesia; guinea pig ileum; omega-conotoxin; pain; neuropathic.  
OS Synthetic.  
OS Conus sp.  
FH Key  
FH Key  
FT Modified\_site 4 Location/Qualifiers  
FT /label= 4Hyp  
FT /note= "4-Hydroxyproline"  
FT Modified\_site 10  
FT /label= 4Hyp  
FT /note= "4-Hydroxyproline"  
FT Modified\_site 21  
FT /label= 4Hyp  
FT /note= "4-Hydroxyproline"  
FT Modified\_site 27  
FT /note= "C-terminal amide"  
PN US5859186-A.  
PD 12-JAN-1999.  
PF 03-JUL-1996; 675354.  
PR 15-APR-1993; US-049794.  
PR 30-DEC-1991; US-814759.  
PR 03-JUL-1996; US-675354.  
PA (NEUR-) NEUREX CORP.  
PI Gohil KC, Justice A, Miljanich GP, Singh T, Valentino KL;  
PI WPI: 99-120002/10.  
DR PT Production of analgesia in mammal - by administration of omega  
PT cono-peptides)  
PS Disclosure: Fig 1A: 59pp; English.  
CC Sequences W95564-573 represent primary sequences of natural omega-  
CC conopeptides. Omega-conopeptides are components of peptide toxins  
CC produced by marine snails of the genus Conus, and which act as calcium  
CC channel blockers. The invention relates to a method of producing  
CC analgesia in a mammal that comprises administering an omega conopeptide  
CC having activities in (a) inhibiting electrically stimulated contraction  
CC of guinea pig ileum and (b) selectively binding to omega conopeptide  
CC MVIIA binding sites in neuronal tissue, where these activities are within  
CC the ranges of those of omega-conotoxins MVIIA and TVIA. The method is  
CC used for treating chronic pain, especially neuropathic pain.  
SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 39; Length 27;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
|||||  
QY 1 YTKRC 5

RESULT 6  
ID W72607 standard; peptide: 27 AA.  
AC W72607;  
DT 06-JAN-1999 (first entry)  
DE Conus genus natural omega-conopeptide GVIA/SNX-124.  
DE Conus genus; marine snail; cone snail; omega-conopeptide; analgesia;  
KW nociceptive pain; neuropathic pain; neuronal tissue; conotoxin;  
KW inflammation; schizophrenia; tardive dyskinesia; acute dystonic reaction;  
KW rheumatoid arthritis; epilepsy.  
OS Conus sp.  
FH Key  
FH Key  
FT Modified\_site 4 Location/Qualifiers  
FT /label= Hyp  
FT /note= "hydroxyproline"  
FT Modified\_site 10

FT /label= Hyp  
FT /note= "hydroxyproline"  
FT 21  
FT /label= Hyp  
FT /note= "hydroxyproline"  
FT  
PN US5824645-A.  
PD 20-OCT-1998.  
PN 01-NOV-1996; 742774.  
PF 15-APR-1993; US-049794.  
PR 30-DEC-1991; US-814759.  
PR 03-JUL-1996; US-675354.  
PR 01-NOV-1996; US-742774.  
PA (NEUR-) NEUREX CORP.  
PI Gohil KC, Justice A, Miljanich GP, Singh T, Valentino KL;  
DR WPI: 98-582596/49.  
PT Treatment of inflammation, comprises administration of  
PT omega-conopeptide - effective to block voltage-gated calcium  
PT channels, bind with high affinity to omega-conopeptide binding site,  
PT and inhibit neuro-transmitter release  
PS Disclosure; Fig 1: 58pp; English.  
CC A method has been developed for the treatment of inflammation in a  
CC subject. The method comprises administration of an omega-conopeptide  
CC effective to: (i) block voltage-gated calcium channels; (ii) bind with  
CC high affinity to an omega-conopeptide binding site; and (iii) inhibit  
CC neurotransmitter release from nervous tissue. The method is used to  
CC treat inflammation and associated pain. The treatment can also be used  
CC to produce analgesia (especially in subjects experiencing neuropathic  
CC pain); and to treat schizophrenia, tardive dyskinesia and acute dystonic  
CC reactions, rheumatoid arthritis, and epilepsy. The present sequence  
CC represents a natural omega-conopeptide. Omega-conopeptides are  
CC components of peptide toxins produced by marine snails of the genus  
CC Conus, and which act as calcium channel blockers.  
SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 37; Length 27;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
|  
|  
|  
|  
Qy 1 YTKRC 5

RESULT 7  
ID R51035 standard; peptide; 27 AA.  
AC R51035;  
DT 24-NOV-1994 (first entry)  
DE N-type calcium channel affinity peptide.  
KW Affinity; N-type; calcium channel; isolation; purification.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT disulfide\_bond 1..16  
FT modified\_site 4  
FT /label= 4Hyp  
FT disulfide\_bond 8..19  
FT modified\_site 10  
FT /label= 4Hyp  
FT disulfide\_bond 15..26  
FT modified\_site 21  
FT /label= 4Hyp  
FT modified\_site 27  
FT /note= "Amidated C-terminal"

J05080696-A.  
PN 22-MAR-1994. 255424.  
PD 01-SEP-1992; JP-255424.  
PR 01-SEP-1992; JP-255424.  
PA (MITU ) MITSUBISHI KASBI CORP.  
DR WPI: 94-132043/16.  
PT Peptide with affinity for N-type calcium channel - useful as  
PT agent for isolation and purification of calcium channel  
PS Claim 1; Page 2; 4pp; Japanese.  
PC This sequence represents a peptide which has affinity for an N-type  
CC calcium channel. This peptide is preferably prepared by standard

CC solid-phase synthesis techniques and is useful as an agent for  
CC isolation and purification of the calcium channel.  
SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 10; Length 27;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
|  
|  
|  
|  
Qy 1 YTKRC 5

RESULT 8  
ID W12959 standard; peptide; 27 AA.  
AC W12959;  
DT 22-APR-1997 (first entry)  
DE Omega conopeptide SNX-124.  
KW Omega conopeptide; analgesic; treatment; neuropathic pain;  
KW inhibition; neuronal damage; schizophrenia; tardive dyskinesia;  
KW analgesia; acute dystonic reactions; inflammation; epilepsy.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT modified\_site 4  
FT /label= Hyp  
FT modified\_site 10  
FT /label= Hyp  
FT modified\_site 21  
FT /label= Hyp  
FT US587454-A.  
PN 24-DEC-1996.  
PD 30-DEC-1991; 814759.  
PR 30-DEC-1991; US-814759.  
PR 30-DEC-1992; WO-U11349.  
PR 15-APR-1993; US-049794.  
PA (NEUR-) NEUREX CORP.  
PI Gohil KC, Justice A, Miljanich GP, Singh T, Valentino KL;  
DR WPI: 97-064830/06.  
PT Omega conopeptide(s) - useful as analgesics, esp. for treating  
PT neuropathic pain  
PS Disclosure; Columns 41-42; 58pp; English.  
CC The present peptide is an omega conopeptide, useful as an  
CC analgesic, especially for treating neuropathic pain. The peptide,  
CC which can be prepared by solid phase synthesis, can also be used to  
CC inhibit neuronal damage and treat schizophrenia, tardive  
CC dyskinesia, acute dystonic reactions, inflammation and epilepsy.  
SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 20; Length 27;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
|  
|  
|  
|  
Qy 1 YTKRC 5

RESULT 9  
ID W19546 standard; peptide; 27 AA.  
AC W19546;  
DT 10-OCT-1997 (first entry)  
DE Natural omega-conopeptide GVIA/SNX-124 used for pain relief.  
KW Conopeptide; cone snail; pain; analgesic; neuropathy; epidural;  
KW N-type voltage-sensitive calcium channel; block; Conus.  
OS Conus sp.  
FH Key Location/Qualifiers  
FT misc\_difference 4  
FT /label= 4Hyp  
FT misc\_difference 10  
FT /label= 4Hyp  
FT misc\_difference 21  
FT /label= 4Hyp  
PN W09701351-A1.



PD 16-JAN-1997.  
 PF 26-JUN-1996: U11041.  
 PR 27-JUN-1995: US-496847.  
 PR 08-MAR-1996: US-613400.  
 PA (NEUR-) NEUREX CORP.  
 PI Adriaenssens PI, Amstutz GA, Bowersox SS, Gadbois T;  
 PI Gohil K, Kristipati R, Luther RR, Pettus MR;  
 PI WPI: 97-100012/09.  
 DR Stable omega conopeptide compositions - for producing analgesia and  
 PT for inhibiting progression of neuropathic pain disorders  
 PS Disclosure: Fig 1: 47pp; English.  
 CC W19544-W19553 are naturally occurring omega conopeptides (OCs) isolated  
 CC from Conus sp. (cone snails). The peptides and their analogues are  
 CC used as analgesics acting by blocking N-type voltage-sensitive calcium  
 CC channels. The OCs can be used to treat neuropathic pain as a result of  
 CC e.g. insult to the spinal cord or peripheral nerves, cancer, bone  
 CC degenerative diseases, AIDS, reflex sympathetic dystrophy, herpes zoster  
 CC neuropathy, diabetic neuropathy, hyperesthesia, allodynia or  
 CC hyperalgesia. The OCs are preferably administered in a medicament via  
 CC an epidural route in a continuous infusion or sustained release  
 CC formulation. The OCs can provide pain relief when administered epidurally  
 CC in the absence of a permeation enhancer, at doses that are comparable to  
 CC effective analgesic doses using intrathecal administration. OC  
 CC formulations comprising an OC and a carboxylic acid buffer anti-oxidant.  
 CC They also confer stability to solutions containing them for prolonged  
 CC treatment methods and long-term storage.  
 SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 23; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
 QY 1 YTKRC 5  
 |||||

RESULT 10  
 ID R37754 standard; peptide; 27 AA.  
 AC R37754;  
 DT 08-SEP-1993 (first entry)  
 DE GVIA/SNX-124.  
 KW Ischaemia; neuronal; omega-conotoxin; OCT; MWIIA; MWIIC; MWIID;  
 KW MWIIB; GVIA; GVIIA; RVIA; SVIA; TVIA; SVIB; SNX-207; stroke;  
 KW delayed treatment; antihistamine; blood pressure;  
 KW N-type voltage-gated Ca currents;  
 KW N-channel mediated neurotransmitter release.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT modified\_site 4  
 FT /note= "hydroxyproline"  
 FT modified\_site 10  
 FT /note= "hydroxyproline"  
 FT modified\_site 21  
 FT /note= "hydroxyproline"  
 FT disulfide\_bond 1..16  
 FT disulfide\_bond 8..19  
 FT disulfide\_bond 15..26  
 PN WO9310145-A.  
 PD 27-MAY-1993.  
 PF 12-NOV-1992: 009786.  
 PR 12-NOV-1991: US-789913.  
 PR 17-JUL-1992: US-916478.  
 PA (NEUR-) NEUREX CORP.  
 PI Bitner RS, Bowersox SS, Fox JA, Miljanich GP, Valentino KL;  
 PI Yamashiro DH;  
 DR WPI: 93-182487/22.  
 PT Redn. of neuronal damage caused by ischaemia - by admin. of epds.  
 PT that bind specifically to omega-conotoxin MWIIA binding sites  
 PS Disclosure: Fig 1: 103pp; English.  
 CC Ischaemia-related neuronal damage in mammals is reduced by admin.,  
 CC 4-24 hr after onset of ischaemia, of a cpd. (I) which binds  
 CC selectively to an omega-conotoxin (OCT) MWIIA site in neuronal

CC tissue. (I) has selectivity at least 100 expressed as ratio of  
 CC binding affinity for the MWIIA site to that for the MWIIC site.  
 CC (I) is one of the OCTs MWIIA, MWIIB, GVIA, GVIIA or RVIA or it is  
 CC the cpd. SNX-207. (I) is esp. used to reduce neuronal damage  
 CC caused by stroke. By delaying admin. for some time (compare  
 CC US051403 where cpds. are given within 1 hr of the onset of  
 CC ischaemia) a greater redn. in neuronal damage is achieved. (I) is  
 CC admin. e.g. by intracerebroventricular (ICV) injection at 0.1-20  
 CC microg/kg, but can also be given i.v. (opt. after treatment with  
 CC antihistamines to minimise redn. in blood pressure caused by (I)).  
 CC (I) is also at least as effective as the specified conotoxins for (I).  
 CC selective inhibition of N-type voltage-gated Ca currents in neuronal  
 CC tissue and (2) selective inhibition of N-channel mediated  
 CC neurotransmitter release in neuronal tissue.  
 CC Primary sequences of omega-conopeptides are given in R37752-62.  
 CC Several analog omega-conopeptides are given in R37763-76.  
 SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 7; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
 QY 1 YTKRC 5  
 |||||

RESULT 11  
 ID R39610 standard; peptide; 27 AA.  
 AC R39610;  
 DT 20-DEC-1993 (first entry)  
 DE GVIA/SNX124.  
 KW Omega conopeptide; OCT; analgesia; inhibition; voltage-gated;  
 KW calcium channel; neurone; contraction; guinea pig; ileum;  
 KW MWIIA; binding site; toxin; marine; snail; Conus; opiod;  
 KW chronic pain; narcotics.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..16  
 FT modified\_site 4  
 FT /note= "4Hyp"  
 FT modified\_site 10  
 FT /note= "4Hyp"  
 FT disulfide\_bond 8..19  
 FT disulfide\_bond 15..26  
 FT modified\_site 21  
 FT /note= "4Hyp"  
 PN WO9313128-A.  
 PD 08-JUL-1993.  
 PF 30-DEC-1992: U11349.  
 PR 30-DEC-1991: US-814759.  
 PA (NEUR-) NEUREX CORP.  
 PI Gohil K, Justice A, Miljanich GP, Singh T, Valentino KL;  
 DR WPI: 93-227270/28.  
 PT Use of omega-cono-peptide(s) which selectively inhibit  
 PT voltage-gated calcium channels - to induce analgesia, enhance  
 PT opiate analgesics, treat pain etc.  
 PS Claim 1; Fig 1: 90pp; English.  
 CC The sequences given in R39608-30 are omega conopeptides (OCTs) and  
 CC derivatives of these, which may be used to produce analgesia in a  
 CC mammal. These OCTs inhibit voltage-gated calcium channels  
 CC selectively in neuronal tissue. This is shown by the peptides  
 CC ability to stimulate contraction in guinea pig ileum and to bind to  
 CC OCT MWIIA binding sites present in neuronal tissue. OCTs are  
 CC components of peptide toxins derived from marine snails of the genus  
 CC Conus, and act as calcium channel blockers. These OCTs may be used  
 CC to replace opiods in the treatment of chronic pain or to reduce the  
 CC opiod dosage required. This helps to reduce dependence on and  
 CC tolerance to opiod narcotics.  
 SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 8; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.49e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
 QY 1 YTKRC 5

RESULT 12

ID R76091 standard; peptide; 27 AA.  
 AC R76091:  
 DT 02-FEB-1996 (first entry)  
 DE Omega conotoxin GVIA peptide.  
 KW Omega conotoxin; marine snail; Conus; voltage-gated Ca channel blocker;  
 KW synaptosome; membrane; fish electric organ; mammalian brain; ischaemia;  
 KW binding protein; binding affinity; stroke.  
 OS Conus sp.  
 FH Key Location/Qualifiers  
 FT disulfide\_bond 1..16  
 FT modified\_site 4  
 FT /label= 4-Hyp  
 FT disulfide\_bond 8..19  
 FT modified\_site 10  
 FT /label= 4-Hyp  
 FT disulfide\_bond 15..26  
 FT modified\_site 21  
 FT /label= 4-Hyp  
 FT modified\_site 27  
 FT /note= "amidated C-terminus"  
 PN US5424218-A.  
 PD 13-JUN-1995.  
 PF 22-NOV-1989; 440094.  
 PR 22-NOV-1989; US-440094.  
 PR 02-AUG-1990; US-561766.  
 PR 23-MAR-1992; US-855269.  
 PR 04-NOV-1993; US-147714.  
 PA (NEUR-) NEUREX CORP.  
 PI Bitner RS, Bowersox SS, Fox JA, Miljanich GP, Valentino KL;  
 PI Yamashiro DH;  
 DR WPI: 95-223694/29.  
 PT Identifying cpds. able to reduce neuronal damage caused by ischaemia  
 PT and ability e.g. to inhibit voltage gated calcium channels  
 PS Disclosure: Fig 1; 3pp; English.  
 CC The peptides R76089-95 are naturally occurring omega conotoxin (OCT)  
 CC peptides derived from marine snails of the Conus genus. The peptide  
 CC sequences were used to chemically synthesise the OCT peptide fragments  
 CC R76096-R76109. The OCT peptides act as voltage-gated Ca channel blockers  
 CC by binding to a 210 kD protein from synaptosomal membrane preparations  
 CC from fish electric organ or mammalian brains. The peptides and their  
 CC synthesised fragments can be used to screen for compounds that bind to  
 CC the OCT binding protein, by displacing a high affinity labelled OCT, such  
 CC as MVIIA, from a synaptosomal membrane preparation. The compounds should  
 CC have binding affinities and activities at least equal to those of the  
 CC natural peptides (K<sub>i</sub> 0.44-324 nM). The screened compounds are  
 CC potentially useful in treating ischaemic conditions, esp. stroke, and can  
 CC reduce sec. anatomical and functional damage associated with those  
 CC conditions.  
 SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 14; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.49e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
 QY 1 YTKRC 5

RESULT 13

ID R32779 standard; peptide; 27 AA.  
 AC R32779:  
 DT 28-JUN-1993 (first entry)  
 DE GVIA omega conotoxin peptide.

KW OCT; neuronal damage reduction; ischemia; secondary damage; stroke.  
 OS Synthetic.  
 PN US5189020-A.  
 PD 23-FEB-1993.  
 PF 02-AUG-1990; 561766.  
 PR 22-NOV-1989; US-440094.  
 PR 02-AUG-1990; US-561766.  
 PA (NEUR-) NEUREX CORP.  
 PI Bitner RS, Bowersox SS, Fox JA, Miljanich GP, Tsubokawa M,  
 PI Valentino KL, Yamashiro DH;  
 DR WPI: 93-085564/10.  
 DT Reducing neuronal damage due to ischaemia - involves using omega  
 PT conotoxin peptide or fragment  
 PS Disclosure: Fig 1; 32pp; English.  
 CC The sequence is that of the GVIA omega conotoxin (OCT) peptide  
 CC which can bind to an OCT binding protein, inhibit voltage-gated  
 CC calcium currents selectively in neuronal tissue and inhibit neuronal  
 CC transmitter release selectively in neuronal tissue. These properties  
 CC all occur within the range of those of MVIIA, GVIIA, RVIIA, or pref.  
 CC MVIIA and GVIA OCTs. The peptide can be used in reducing or  
 CC preventing both anatomical and functional secondary damage related  
 CC to ischemia, generally as associated with stroke.  
 SQ Sequence 27 AA;

Query Match 20.8%; Score 5; DB 6; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 2.49e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ytkrc 26  
 QY 1 YTKRC 5

RESULT 14

ID W32240 standard; Protein; 47 AA.  
 AC W32240:  
 DT 20-APR-1998 (first entry)  
 DE B chain sequence of a single chain monellin analogue.  
 KW Monellin; analogue; thermostable; sweetener; food; beverage;  
 KW linker; B chain; A chain.  
 OS Synthetic.  
 PN US5672372-A.  
 PD 30-SEP-1997.  
 PF 06-JUN-1995; 469854.  
 PR 30-MAR-1990; US-502257.  
 PR 19-JUN-1987; US-064341.  
 PR 19-JUN-1987; US-064343.  
 PR 04-NOV-1987; US-117124.  
 PR 18-JAN-1990; US-465585.  
 PR 02-NOV-1993; US-146326.  
 PR 06-JUN-1995; US-469854.  
 PA (LUCK-) LUCKY BIOTECH CORP.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Cho JM, Kim S;  
 DR WPI: 97-488831/45.  
 PT Sweetening food compositions - with thermostable single chain  
 PT monellin analogues  
 PS Claim 2: Column 13; 10pp; English.  
 CC This is the B chain sequence used in the construction of a single chain  
 CC monellin analogue. This consists of peptide residues 1-46 of the B chain  
 CC of native monellin. This can be used in a method for sweetening a food  
 CC composition. The method comprises including in the composition the  
 CC single chain sweet proteinaceous analogue of monellin. The monellin  
 CC analogue contains this B chain and residues 6-45 of the A chain of native  
 CC monellin modified only by conservative substitution. The B chain and the  
 CC A chain are linked by a covalent bond or a hydrophilic, physiologically  
 CC acceptable covalent linker capable of providing a spacing length  
 CC equivalent to a peptide of 1-10 amino acids, selected so as to reside on  
 CC the external portion of the molecule and not to disturb the native  
 CC conformation. This protein can be used for sweetening foods and  
 CC beverages. The monellin analogues are thermostable and retain  
 CC high-intensity sweetening power even after heating to 100 degree. C at  
 CC acidic pH.

SQ Sequence 47 AA;

Query Match 20.8%; Score 5; DB 27; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 tfnk 38  
|||||  
QY 6 TFNKV 10

RESULT 15

ID R88034 standard; Peptide; 47 AA.  
AC R88034;  
DT 10-JUL-1996 (first entry)  
DE B chain sequence of monellin (residues 1-46).  
KW monellin; analogue; Serendipity Berry; Dioscoreophyllum comminisii;  
KW sweetener; food; beverage; single chain form.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT peptide 2..47  
FT /note= "residues 1-46 of native B chain"  
PN US5478923-A.  
PD 26-DEC-1995.  
PF 19-JUN-1987; 064341.  
PR 19-JUN-1987; US-064341.  
PR 04-NOV-1987; US-117124.  
PR 18-JAN-1990; US-465385.  
PR 30-MAR-1990; US-502257.  
PR 25-MAR-1994; US-218264.  
PA (LUCK-) LUCKY CO LTD.  
PA (REGC) UNIV CALIFORNIA.  
PI Cho JM, Kim S;  
DR WPI; 96-057806/06.  
PT New single chain analogues of monellin - useful as sweeteners in  
PT foods, beverages, pharmaceuticals etc., are stable at high temps.  
PT and acidic pH  
PS Claim 2; Column 13; 10pp: English.  
CC The present sequence represents residues 1-46 of the native B chain of  
CC monellin preceded by a methionine. Monellin is isolated from  
CC "Serendipity Berries" of the West African plant Dioscoreophyllum  
CC comminisii. The protein can be used as a substitute for sugar in  
CC sweetening foods and beverages. Single chain forms of monellin, which  
CC are capable of maintaining their 3-D conformation under conditions which  
CC would ordinarily denature the native protein and retain their sweet  
CC taste, are provided in the invention. (See R88035-49).  
SQ Sequence 47 AA;

Query Match 20.8%; Score 5; DB 17; Length 47;  
Best Local Similarity 100.0%; Pred. No. 2.49e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 tfnk 38  
|||||  
QY 6 TFNKV 10

Search completed: Sat Aug 28 14:46:07 1999  
Job time : 69 secs.

**This Page Blank (uspto)**

\*\*\*\*\*  
WORLD  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:43:57 1999; MasPar time 4.91 Seconds  
Tabular output not generated.  
195.790 Million cell updates/sec

Title: >US-09-049-696-43  
Description: (1-24) from US09049696.ppe  
Perfect Score: 24  
Sequence: 1 YTKRCTFNKVTGLYEKCEFLVQS 24

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 2.562; Variance 0.399; scale 6.421

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	25.0	160	2	H69804 transcription regulat	2.60e+00
2	6	25.0	279	2	T02471 hypothetical protein	2.60e+00
3	6	25.0	293	2	B70319 conserved hypothetical	2.60e+00
4	6	25.0	341	2	I61725 natural killer associ	2.60e+00
5	6	25.0	385	2	A47701 aspartic proteinase A	2.60e+00
6	6	25.0	435	2	G01923 natural killer associ	2.60e+00
7	6	25.0	452	2	G01923 KIR (cl-5) NK recepto	2.60e+00
8	6	25.0	544	2	S65231 tRNA-pseudouridine sy	2.60e+00
9	6	25.0	585	2	S48951 hypothetical protein	2.60e+00
10	6	25.0	1450	2	A44027 165K myofibrillar M-b	2.60e+00
11	6	25.0	2696	1	I38344 titin, cardiac muscle	2.60e+00
12	5	20.8	27	2	S15030 transferrin-binding p	1.08e+02
13	5	20.8	49	2	G70152 hypothetical protein	1.08e+02
14	5	20.8	51	1	MLDRB monellin chain B - se	1.08e+02
15	5	20.8	73	1	NTKN6G omega-conotoxin GVIB	1.08e+02
16	5	20.8	78	2	D45357 orf A immediately dow	1.08e+02
17	5	20.8	91	2	A49754 hypothetical protein	1.08e+02
18	5	20.8	96	2	A25274 phosphodiesterase I (	1.08e+02
19	5	20.8	101	2	S59127 trp-like protein - Af	1.08e+02
20	5	20.8	103	2	I52722 gene TCFA protein - h	1.08e+02
21	5	20.8	104	2	B69058 hypothetical protein	1.08e+02
22	5	20.8	105	2	A69329 hypothetical protein	1.08e+02
23	5	20.8	119	2	A69361 conserved hypothetical	1.08e+02

24	20.8	123	2	F64342	hypothetical protein	1.08e+02
25	20.8	130	2	JC4671	ribosomal protein S24	1.08e+02
26	20.8	130	2	H70010	polyribonucleotide nu	1.08e+02
27	20.8	131	2	H70358	hypothetical protein	1.08e+02
28	20.8	131	2	S40161	ribosomal protein S24	1.08e+02
29	20.8	132	1	R3XL19	ribosomal protein S24	1.08e+02
30	20.8	133	2	B70027	hypothetical protein	1.08e+02
31	20.8	133	1	R3HI24	ribosomal protein S24	1.08e+02
32	20.8	133	1	R3RT24	ribosomal protein S24	1.08e+02
33	20.8	133	2	JH0213	ribosomal protein S24	1.08e+02
34	20.8	139	2	S25970	hypothetical protein	1.08e+02
35	20.8	144	2	S32235	ribosomal protein L11	1.08e+02
36	20.8	144	2	B41773	ribosomal protein L11	1.08e+02
37	20.8	144	2	S54718	ribosomal protein L11	1.08e+02
38	20.8	144	2	S40772	ribosomal protein L11	1.08e+02
39	20.8	147	2	JE0148	obesity protein - hum	1.08e+02
40	20.8	147	2	F71802	hypothetical protein	1.08e+02
41	20.8	147	2	F64715	hypothetical protein	1.08e+02
42	20.8	160	2	B71238	hypothetical protein	1.08e+02
43	20.8	162	2	S17803	hypothetical protein	1.08e+02
44	20.8	162	2	PN0556	phytochrome - wild ri	1.08e+02
45	20.8	166	2	I53166	leptin precursor - hu	1.08e+02
46	20.8	166	2	PN0555	phytochrome - wild ri	1.08e+02
47	20.8	167	2	B69687	menaquinol:cytochrome	1.08e+02
48	20.8	167	1	LTHU	leptin precursor - hu	1.08e+02
49	20.8	169	2	PN0360	phytochrome - long-st	1.08e+02
50	20.8	170	2	PN0557	phytochrome - wild ri	1.08e+02
51	20.8	171	2	PN0553	phytochrome - wild ri	1.08e+02
52	20.8	171	2	A46618	collagen-induced plat	1.08e+02
53	20.8	171	2	G69077	feritin like protein	1.08e+02
54	20.8	171	2	PN0558	phytochrome - wild ri	1.08e+02
55	20.8	172	2	JC4192	ribosomal protein L11	1.08e+02
56	20.8	172	2	PN0552	phytochrome - wild ri	1.08e+02
57	20.8	173	2	E64540	hypothetical protein	1.08e+02
58	20.8	173	2	I40471	hypothetical protein	1.08e+02
59	20.8	176	2	PN0559	phytochrome - wild ri	1.08e+02
60	20.8	177	1	OVF0P	olfactory protein-pre	1.08e+02
61	20.8	177	2	S45494	isp3 protein - fissio	1.08e+02
62	20.8	180	2	I55371	ADP-ribosylation fact	1.08e+02
63	20.8	180	2	JC4948	ADP-ribosylation fact	1.08e+02
64	20.8	180	2	G69974	methyltransferase hom	1.08e+02
65	20.8	180	2	B38622	ADP-ribosylation fact	1.08e+02
66	20.8	184	2	S72231	transcription factor	1.08e+02
67	20.8	186	2	E70152	3-methyladenine DNA g	1.08e+02
68	20.8	190	2	A38256	GTP cyclohydrolase I	1.08e+02
69	20.8	195	2	H71266	hypothetical protein	1.08e+02
70	20.8	201	1	Q0B5I8	UL92 protein - human	1.08e+02
71	20.8	201	2	G64013	hypothetical protein	1.08e+02
72	20.8	210	2	C36864	probable glucose-1-ph	1.08e+02
73	20.8	212	2	S71114	ribosomal protein S19	1.08e+02
74	20.8	219	2	E69898	general secretion pat	1.08e+02
75	20.8	220	2	C69824	phosphoglycolate phos	1.08e+02
76	20.8	221	2	B64155	hypothetical protein	1.08e+02
77	20.8	224	2	F69321	conserved hypothetical	1.08e+02
78	20.8	227	2	F69896	conserved hypothetical	1.08e+02
79	20.8	228	2	D69911	hypothetical protein	1.08e+02
80	20.8	228	2	B69777	conserved hypothetical	1.08e+02
81	20.8	228	2	C69859	two-component respons	1.08e+02
82	20.8	231	2	S50462	hypothetical protein	1.08e+02
83	20.8	231	2	S75380	hypothetical protein	1.08e+02
84	20.8	234	2	S06141	ATP-dependent DNA hel	1.08e+02
85	20.8	235	2	D71808	hypothetical protein	1.08e+02
86	20.8	239	2	G02630	Fcalphaarb - human	1.08e+02
87	20.8	241	2	B71237	hypothetical protein	1.08e+02
88	20.8	244	2	E65128	retrovirus-related re	1.08e+02
89	20.8	245	2	F44490	transcription factor	1.08e+02
90	20.8	249	2	A45130	ATP-dependent RNA hel	1.08e+02
91	20.8	249	2	S18677	rap30 protein - rat	1.08e+02
92	20.8	249	2	S35491	conserved hypothetical	1.08e+02
93	20.8	254	2	C69787	hypothetical protein,	1.08e+02
94	20.8	255	2	S07553	beta-etherase ligf -	1.08e+02
95	20.8	257	2	S33314	NH(3)-dependent NAD+	1.08e+02
96	20.8	259	2	G64468		

97	5	20.8	262	2	A70348	1.08e+02	hypothetical protein	170	5	20.8	420	2	S21052	interleukin-5 recepto	1.08e+02
98	5	20.8	262	2	S76610	1.08e+02	hypothetical protein	171	5	20.8	422	1	A44502	nucleoprotein - Chand	1.08e+02
99	5	20.8	263	2	S73129	1.08e+02	hypothetical protein	172	5	20.8	422	1	JGAGLR	lactose-binding prote	1.08e+02
100	5	20.8	265	1	ORECFH	1.08e+02	ferrichrome transport	173	5	20.8	425	2	G70848	probable glyA2 protei	1.08e+02
101	5	20.8	274	2	S11489	1.08e+02	formamidopyrimidine-D	174	5	20.8	426	2	I48379	gene hb protein - mou	1.08e+02
102	5	20.8	279	2	E64109	1.08e+02	dimethylsulfoxide red	175	5	20.8	427	2	S75210	glycine hydroxymethyl	1.08e+02
103	5	20.8	285	2	E70151	1.08e+02	chemotaxis protein me	176	5	20.8	429	2	C64928	fixc protein homolog	1.08e+02
104	5	20.8	285	2	H69802	1.08e+02	conserved hypothetical	177	5	20.8	433	1	E64242	GTP-binding protein o	1.08e+02
105	5	20.8	287	2	JH0332	1.08e+02	IgA (Fc) receptor, my	178	5	20.8	439	2	JCS253	TEA domain-containing	1.08e+02
106	5	20.8	288	2	S47046	1.08e+02	glucose-1-phosphate t	179	5	20.8	440	2	A42136	transcription factor	1.08e+02
107	5	20.8	289	2	C69349	1.08e+02	conserved hypothetical	180	5	20.8	442	2	A71969	probable histidine ki	1.08e+02
108	5	20.8	289	2	S33294	1.08e+02	probable glucose-1-ph	181	5	20.8	445	2	A48881	serotonin receptor 7	1.08e+02
109	5	20.8	290	2	E69366	1.08e+02	branched-chain amino	182	5	20.8	446	2	S60180	branched-chain amino	1.08e+02
110	5	20.8	290	2	T00103	1.08e+02	glucose-1-phosphate t	183	5	20.8	448	2	A47519	serotonin receptor 7	1.08e+02
111	5	20.8	290	2	S41536	1.08e+02	glucose-1-phosphate t	184	5	20.8	448	2	S40687	serotonin receptor 7	1.08e+02
112	5	20.8	291	2	S05508	1.08e+02	photosystem II oxygen	185	5	20.8	448	2	S36402	serotonin receptor 7	1.08e+02
113	5	20.8	292	2	S23342	1.08e+02	glucose-1-phosphate t	186	5	20.8	452	2	E71154	hypothetical protein	1.08e+02
114	5	20.8	292	2	S15301	1.08e+02	glucose-1-phosphate t	187	5	20.8	453	2	H71732	cytochrome D ubiquino	1.08e+02
115	5	20.8	293	2	H65182	1.08e+02	glucose-1-phosphate t	188	5	20.8	457	2	A36329	hypothetical protein	1.08e+02
116	5	20.8	293	2	F64969	1.08e+02	glucose-1-phosphate t	189	5	20.8	459	2	T02569	probable myrosinase-b	1.08e+02
117	5	20.8	293	2	S78544	1.08e+02	glucose-1-phosphate t	190	5	20.8	464	2	C64462	adenosylmethionine--8	1.08e+02
118	5	20.8	295	2	A49906	1.08e+02	glucose-1-phosphate t	191	5	20.8	471	2	S28476	rflB protein - Vibrio,	1.08e+02
119	5	20.8	295	2	S46718	1.08e+02	hypothetical protein	192	5	20.8	474	2	S24601	hypothetical protein,	1.08e+02
120	5	20.8	298	2	A44878	1.08e+02	protein kinase (EC 2.	193	5	20.8	475	2	A30610	chloroplast rbc L gen	1.08e+02
121	5	20.8	299	2	A45720	1.08e+02	p35 - Bombyx mori nuc	194	5	20.8	477	2	H64210	eggshell protein p48	1.08e+02
122	5	20.8	299	1	WMNV35	1.08e+02	early 35K protein - A	195	5	20.8	480	2	JN0558	UMP synthase - bovine	1.08e+02
123	5	20.8	300	2	A41179	1.08e+02	protein kinase PC-1 (	196	5	20.8	480	2	A30148	UMP synthase - human	1.08e+02
124	5	20.8	305	1	S71284	1.08e+02	myb-related protein,	197	5	20.8	480	2	B64308	hypothetical protein	1.08e+02
125	5	20.8	310	2	S58090	1.08e+02	probable membrane pro	198	5	20.8	487	2	S41951	UTP-glucose glucosyl	1.08e+02
126	5	20.8	319	2	H70182	1.08e+02	hypothetical protein	199	5	20.8	495	1	MXRSA	nonstructural protein	1.08e+02
127	5	20.8	319	2	S48811	1.08e+02	1-aminocyclopropane-1	200	5	20.8	503	2	S31940	starch-degrading enzy	1.08e+02
128	5	20.8	320	2	S22488	1.08e+02	ethylene-forming enzy	201	5	20.8	508	1	QJNT6A	photosystem II chloro	1.08e+02
129	5	20.8	320	1	KIECFA	1.08e+02	6-phosphofructokinase	202	5	20.8	508	1	QJZMBB	photosystem II chloro	1.08e+02
130	5	20.8	320	1	RREC	1.08e+02	RNA-directed DNA poly	203	5	20.8	508	1	QJSP6A	photosystem II chloro	1.08e+02
131	5	20.8	324	2	S74521	1.08e+02	hypothetical protein	204	5	20.8	508	1	QJRP6A	photosystem II chloro	1.08e+02
132	5	20.8	324	2	S75735	1.08e+02	probable malate dehyd	205	5	20.8	508	2	S12129	photosystem II chloro	1.08e+02
133	5	20.8	326	2	D70622	1.08e+02	probable prsA protein	206	5	20.8	508	2	S14140	photosystem II chloro	1.08e+02
134	5	20.8	328	2	S78498	1.08e+02	probable glucose-1-ph	207	5	20.8	508	2	S12132	photosystem II chloro	1.08e+02
135	5	20.8	330	2	B64688	1.08e+02	glyceraldehyde-3-phos	208	5	20.8	508	2	S01385	photosystem II chloro	1.08e+02
136	5	20.8	330	2	I46691	1.08e+02	CD86 precursor - rabb	209	5	20.8	508	2	S04100	photosystem II chloro	1.08e+02
137	5	20.8	333	2	S21053	1.08e+02	interleukin-5 recepto	210	5	20.8	514	2	C64154	hypothetical protein	1.08e+02
138	5	20.8	335	2	A40267	1.08e+02	interleukin-5 recepto	211	5	20.8	517	2	S65305	probable membrane pro	1.08e+02
139	5	20.8	336	2	C64468	1.08e+02	hypothetical protein	212	5	20.8	522	2	G05222	sorting nexin 1 - hum	1.08e+02
140	5	20.8	336	2	E69214	1.08e+02	GTP-binding protein,	213	5	20.8	540	2	A70358	topoisomerase I - Aqu	1.08e+02
141	5	20.8	337	2	F69860	1.08e+02	dolichol phosphate ma	214	5	20.8	542	2	T02379	hypothetical protein	1.08e+02
142	5	20.8	341	2	S23907	1.08e+02	translation elongatio	215	5	20.8	545	2	F64181	CYP synthase (EC 6.3.	1.08e+02
143	5	20.8	343	2	H69018	1.08e+02	formate dehydrogenase	216	5	20.8	551	2	S60994	VPS17 protein - yeast	1.08e+02
144	5	20.8	348	2	A56247	1.08e+02	natural killer cell-a	217	5	20.8	553	2	S15809	phosphoprotein phosph	1.08e+02
145	5	20.8	356	2	E71252	1.08e+02	probable glycerol-3-p	218	5	20.8	555	2	C45868	glycerol-3-phosphate	1.08e+02
146	5	20.8	357	2	F69987	1.08e+02	spore coat protein ho	219	5	20.8	557	2	S61980	hypothetical protein	1.08e+02
147	5	20.8	360	2	T01161	1.08e+02	hypothetical protein	220	5	20.8	568	2	A71656	ribosomal protein S1	1.08e+02
148	5	20.8	361	2	S71182	1.08e+02	probable polygalactur	221	5	20.8	583	2	T02382	hypothetical protein	1.08e+02
149	5	20.8	367	2	S22349	1.08e+02	hypothetical protein	222	5	20.8	585	2	H69857	ABC transporter (ATP-	1.08e+02
150	5	20.8	370	2	B70310	1.08e+02	conserved hypothetical	223	5	20.8	588	2	S30929	catechol oxidase (EC	1.08e+02
151	5	20.8	370	2	S70157	1.08e+02	cpSF protein, 40.6K -	224	5	20.8	588	2	S34786	catechol oxidase (EC	1.08e+02
152	5	20.8	374	2	S30034	1.08e+02	translocating chain-a	225	5	20.8	596	1	JU0345	major capsid protein	1.08e+02
153	5	20.8	381	2	C64896	1.08e+02	hypothetical protein	226	5	20.8	598	2	T02265	hypothetical protein	1.08e+02
154	5	20.8	387	2	C71827	1.08e+02	hypothetical protein	227	5	20.8	600	2	T02266	hypothetical protein	1.08e+02
155	5	20.8	393	2	S48989	1.08e+02	branched-chain-amino	228	5	20.8	604	2	I35917	glycoprotein IIB - ye	1.08e+02
156	5	20.8	395	2	A37350	1.08e+02	cyclin B - starfish (	229	5	20.8	605	2	A27274	riboflavin I precurs	1.08e+02
157	5	20.8	396	2	S21050	1.08e+02	interleukin-5 recepto	230	5	20.8	607	2	A26168	riboflavin I precurs	1.08e+02
158	5	20.8	397	2	E69858	1.08e+02	conserved hypothetical	231	5	20.8	612	2	S73611	glucose inhibited div	1.08e+02
159	5	20.8	398	2	H70393	1.08e+02	hemolysin - Aquifex a	232	5	20.8	613	2	A69535	aldehyde ferredoxin o	1.08e+02
160	5	20.8	399	2	S71927	1.08e+02	ATP-gated ion channel	233	5	20.8	614	2	B43935	spat protein - Bacill	1.08e+02
161	5	20.8	404	2	A47425	1.08e+02	serotonin receptor 7	234	5	20.8	618	2	A36907	pyocin S1 SIA subunit	1.08e+02
162	5	20.8	408	2	S54467	1.08e+02	hypothetical protein	235	5	20.8	619	2	F70480	conserved hypothetical	1.08e+02
163	5	20.8	410	2	A64357	1.08e+02	succinyl-diaminopimel	236	5	20.8	630	2	E71451	hypothetical protein	1.08e+02
164	5	20.8	412	2	G02453	1.08e+02	NN8-4AG - human (frag	237	5	20.8	630	2	G64226	hypothetical protein	1.08e+02
165	5	20.8	413	2	S65948	1.08e+02	hemolin - cecropia mo	238	5	20.8	632	2	S73824	hypothetical protein	1.08e+02
166	5	20.8	413	2	A37778	1.08e+02	hemolin precursor - c	239	5	20.8	641	1	ZZRNQ	modulation protein no	1.08e+02
167	5	20.8	414	2	JQ1016	1.08e+02	glycine hydroxymethyl	240	5	20.8	642	1	I40071	threonine--trNA ligas	1.08e+02
168	5	20.8	414	2	S60190	1.08e+02	vicilin - zamia furfu	241	5	20.8	651	2	G69177	methionine--trNA liga	1.08e+02
169	5	20.8	415	2	G69885	1.08e+02	processing proteinase	242	5	20.8	654	2	E64245	periplasmic phosphate	1.08e+02

243	5	20.8	658	2	A39018	carnitine O-palmitoyl	1.08e+02	316	5	20.8	1505	2	S28079	DNA-directed DNA poly	1.08e+02
244	5	20.8	665	2	E70118	flagellar hook-associ	1.08e+02	317	5	20.8	1589	1	RGBYCS	cell division control	1.08e+02
245	5	20.8	686	2	E71106	probable DNA topoisom	1.08e+02	318	5	20.8	1603	1	BVASA1	arom protein - Emeric	1.08e+02
246	5	20.8	690	2	C36907	pyocin S2 S2A subunit	1.08e+02	319	5	20.8	1607	2	T03022	MAP kinase kinase kin	1.08e+02
247	5	20.8	699	2	S67773	hypothetical elongatio	1.08e+02	320	5	20.8	1615	1	WMTM18	183.3K protein - toba	1.08e+02
248	5	20.8	701	2	S31150	translatation elongatio	1.08e+02	321	5	20.8	1729	2	S57596	ribosomal RNA process	1.08e+02
249	5	20.8	702	1	VCBPAP	coat protein vp1 - Al	1.08e+02	322	5	20.8	1786	2	H71527	probable excinuclease	1.08e+02
250	5	20.8	705	2	A69720	sublancin 168 lantibi	1.08e+02	323	5	20.8	1809	2	S73729	tuberosus sclerosis 2	1.08e+02
251	5	20.8	707	2	T00665	hypothetical protein	1.08e+02	324	5	20.8	1937	2	T03224	probable polyketide s	1.08e+02
252	5	20.8	721	2	JC4643	lanosterol synthase (	1.08e+02	325	5	20.8	1953	2	S63244	BN1 protein - yeast	1.08e+02
253	5	20.8	725	2	A45033	myelin transcription	1.08e+02	326	5	20.8	2109	2	I38414	transcription factor	1.08e+02
254	5	20.8	730	2	S48813	hypothetical protein	1.08e+02	327	5	20.8	2167	2	S50658	bud emergence protein	1.08e+02
255	5	20.8	764	2	F71863	phenylalanyl-tRNA syn	1.08e+02	328	5	20.8	2178	2	S29237	calcium channel prote	1.08e+02
256	5	20.8	764	2	B64570	phenylalanine--tRNA 1	1.08e+02	329	5	20.8	2222	2	A37490	voltage-dependent cal	1.08e+02
257	5	20.8	781	2	T02272	hypothetical protein	1.08e+02	330	5	20.8	2233	2	S63347	acetyl-CoA carboxylas	1.08e+02
258	5	20.8	788	1	JDVLHH	DNA-directed DNA poly	1.08e+02	331	5	20.8	2251	2	B54972	voltage-dependent cal	1.08e+02
259	5	20.8	794	2	T02171	hypothetical protein	1.08e+02	332	5	20.8	2259	2	S29236	calcium channel prote	1.08e+02
260	5	20.8	796	2	G64570	biotin sulfoxide redu	1.08e+02	333	5	20.8	2270	2	A54972	voltage-dependent cal	1.08e+02
261	5	20.8	796	2	H71865	probable s/n-oxide re	1.08e+02	334	5	20.8	2272	2	C54972	voltage-dependent cal	1.08e+02
262	5	20.8	804	2	D69650	leucine--tRNA ligase	1.08e+02	335	5	20.8	2324	2	A29924	acetyl-CoA carboxylas	1.08e+02
263	5	20.8	810	2	S44653	f42h10.5 protein - Ca	1.08e+02	336	5	20.8	2339	2	S41121	acetyl-CoA carboxylas	1.08e+02
264	5	20.8	817	2	A48721	titin, muscle - chick	1.08e+02	337	5	20.8	2340	2	B71704	cell surface antigen	1.08e+02
265	5	20.8	819	2	C71544	probable leucyl tRNA	1.08e+02	338	5	20.8	2345	2	A35578	acetyl-CoA carboxylas	1.08e+02
266	5	20.8	820	2	T02152	Lu-ECAM-1 protein - b	1.08e+02	339	5	20.8	2346	2	I38928	acetyl-CoA carboxylas	1.08e+02
267	5	20.8	852	2	D71896	aconitate hydratase -	1.08e+02	340	5	20.8	3084	1	MMMSA	laminin alpha-1 chain	1.08e+02
268	5	20.8	853	2	A71339	probable outer membra	1.08e+02	341	5	20.8	3125	1	GNVSPP	genome polyprotein -	1.08e+02
269	5	20.8	853	2	C64617	aconitate hydratase (	1.08e+02	342	5	20.8	3140	1	GNVSRA	genome polyprotein -	1.08e+02
270	5	20.8	859	2	S62441	chitin synthase (EC 2	1.08e+02	343	5	20.8	3140	2	A47508	genome polyprotein -	1.08e+02
271	5	20.8	859	2	T01461	hypothetical protein	1.08e+02	344	5	20.8	3141	1	GNVSPD	genome polyprotein -	1.08e+02
272	5	20.8	861	2	T00434	kinesin heavy chain h	1.08e+02	345	5	20.8	3141	1	GNVSPV	genome polyprotein -	1.08e+02
273	5	20.8	867	2	S72842	methionine synthase m	1.08e+02	346	5	20.8	3461	2	S58870	reelin precursor - mo	1.08e+02
274	5	20.8	869	2	S76720	hypothetical protein	1.08e+02	347	5	20.8	3759	2	A35085	trithorax protein - f	1.08e+02
275	5	20.8	875	2	G69910	hypothetical protein	1.08e+02	348	5	20.8	4092	2	S38128	dynhein heavy chain, c	1.08e+02
276	5	20.8	875	2	A57080	cell surface antigen	1.08e+02	349	5	20.8	6805	2	S20901	titin - rabbit (fragm	1.08e+02
277	5	20.8	878	2	C71305	probable leucyl-tRNA	1.08e+02	350	4	16.7	26	2	C42762	multicatalytic endope	2.60e+03
278	5	20.8	880	2	E14420	valine--tRNA ligase (	1.08e+02	351	4	16.7	33	2	S48211	phospholipase D - boy	2.60e+03
279	5	20.8	900	2	E71623	protein with DnaJ dom	1.08e+02	352	4	16.7	34	2	S65371	tRNA-guanine transgly	2.60e+03
280	5	20.8	905	2	T02205	Lu-ECAM-1 protein - b	1.08e+02	353	4	16.7	35	2	A60496	12K serum protein, be	2.60e+03
281	5	20.8	905	2	A27410	plasma cell membrane	1.08e+02	354	4	16.7	35	2	S49309	oncofetal protein (OF	2.60e+03
282	5	20.8	913	3	JG0168	gob-5 protein - Mouse	1.08e+02	355	4	16.7	55	2	S13005	H+-transporting ATP s	2.60e+03
283	5	20.8	925	2	A39216	plasma cell membrane	1.08e+02	356	4	16.7	59	2	PX0083	glucagon-degrading en	2.60e+03
284	5	20.8	935	2	B64083	hemoglobin-binding pr	1.08e+02	357	4	16.7	60	2	S59092	hypothetical protein	2.60e+03
285	5	20.8	956	2	G70327	isoleucyl-tRNA synthet	1.08e+02	358	4	16.7	61	2	J70366	hypothetical protein	2.60e+03
286	5	20.8	965	2	T02021	P109 protein - silkw	1.08e+02	359	4	16.7	64	2	C69386	conserved hypothetica	2.60e+03
287	5	20.8	986	2	S12021	thermoactive cellulul	1.08e+02	360	4	16.7	64	2	A36689	50K integrin-associat	2.60e+03
288	5	20.8	1000	2	I46521	titin - rabbit (fragm	1.08e+02	361	4	16.7	64	2	S63681	signal transducer and	2.60e+03
289	5	20.8	1017	2	S67804	LRG1 protein - yeast	1.08e+02	362	4	16.7	64	2	A48411	Myf5 homolog - chicke	2.60e+03
290	5	20.8	1037	2	A60163	glycoprotein iib - ra	1.08e+02	363	4	16.7	65	2	S63679	signal transducer and	2.60e+03
291	5	20.8	1039	2	A34269	integrin alpha-2b cha	1.08e+02	364	4	16.7	82	2	I51137	MHC class II beta cha	2.60e+03
292	5	20.8	1042	2	H70203	isoleucine--tRNA liga	1.08e+02	365	4	16.7	82	2	I50984	MHC class II beta cha	2.60e+03
293	5	20.8	1053	2	S44250	integrin alpha-5 chal	1.08e+02	366	4	16.7	82	2	I51148	MHC class II beta cha	2.60e+03
294	5	20.8	1064	2	T01960	hypothetical protein	1.08e+02	367	4	16.7	82	2	I50968	MHC class II beta cha	2.60e+03
295	5	20.8	1066	2	F71100	probable isoleucyl-tR	1.08e+02	368	4	16.7	82	2	I50968	MHC class II beta cha	2.60e+03
296	5	20.8	1073	2	S69079	hypothetical protein	1.08e+02	369	4	16.7	83	2	T03349	gene e18 protein - la	2.60e+03
297	5	20.8	1120	2	T01863	hypothetical protein	1.08e+02	370	4	16.7	91	2	T03202	hypothetical protein	2.60e+03
298	5	20.8	1123	2	D36790	hypothetical protein	1.08e+02	371	4	16.7	92	2	JC2576	hypothetical 10.0K pr	2.60e+03
299	5	20.8	1128	2	S03728	phytochrome (clone ph	1.08e+02	372	4	16.7	94	2	JC5913	hypothetical 10.0K pr	2.60e+03
300	5	20.8	1131	2	JQ0382	phytochrome A - maize	1.08e+02	373	4	16.7	111	2	S29738	pregnancy zone protei	2.60e+03
301	5	20.8	1165	2	A48667	peroxisomal assembly	1.08e+02	374	4	16.7	112	2	S69140	CAMP-binding protein	2.60e+03
302	5	20.8	1192	2	G70513	probable meth - Mycob	1.08e+02	375	4	16.7	114	2	H71019	hypothetical protein	2.60e+03
303	5	20.8	1195	2	S76592	hypothetical protein	1.08e+02	376	4	16.7	115	2	G71058	hypothetical protein	2.60e+03
304	5	20.8	1203	2	S27545	pullulanase - Thermoa	1.08e+02	377	4	16.7	118	2	S51207	cruxhalorhodopsin-1	2.60e+03
305	5	20.8	1226	2	B64479	protoporphyrin IX mag	1.08e+02	378	4	16.7	121	2	A71344	conserved hypothetica	2.60e+03
306	5	20.8	1313	1	VGUPT	M polyprotein - Punta	1.08e+02	379	4	16.7	121	2	A71031	hypothetical protein	2.60e+03
307	5	20.8	1316	2	G70535	probable rpoC protein	1.08e+02	380	4	16.7	122	2	H70370	holo-(acyl-carrier pr	2.60e+03
308	5	20.8	1316	2	S31146	DNA-directed RNA poly	1.08e+02	381	4	16.7	123	2	S34621	hypothetical protein	2.60e+03
309	5	20.8	1322	2	T01842	hypothetical protein	1.08e+02	382	4	16.7	125	2	T03359	gene e8 protein - lac	2.60e+03
310	5	20.8	1398	2	S56814	microtubule-interacti	1.08e+02	383	4	16.7	133	2	I44464	hemiptericin - Pyrrho	2.60e+03
311	5	20.8	1405	1	S13421	polymeric globin alph	1.08e+02	384	4	16.7	133	2	I45927	mb-1 - bovine	2.60e+03
312	5	20.8	1473	2	S53599	DNA topoisomerase (AT	1.08e+02	395	4	16.7	137	2	A58930	succinate dehydrogena	2.60e+03
313	5	20.8	1475	2	A4765	alpha-amylase (EC 3.2	1.08e+02	386	4	16.7	142	2	I51063	MHC class II beta cha	2.60e+03
314	5	20.8	1481	2	A47341	amylopullulanase prec	1.08e+02	387	4	16.7	142	2	I51061	MHC class II beta cha	2.60e+03
315	5	20.8	1481	2	S28669	amylopullulanase prec	1.08e+02	388	4	16.7	143	2	B71372	conserved hypothetica	2.60e+03

389	4	16.7	143	2	D70332	hypothetical protein	2.60e+03	462	2	G70452	hypothetical protein	2.60e+03
390	4	16.7	145	2	S30247	Hin-1 protein - human	2.60e+03	463	2	B43912	myogenic regulatory f	2.60e+03
391	4	16.7	145	2	S68215	Mas 20 protein - huma	2.60e+03	464	4	F47188	MHC class II histocom	2.60e+03
392	4	16.7	147	2	D69339	conserved hypothetical	2.60e+03	465	4	S06553	finger protein (clone	2.60e+03
393	4	16.7	147	2	B70436	hypothetical protein	2.60e+03	466	248	S49323	chymotrypsin (EC 3.4.	2.60e+03
394	4	16.7	152	2	C39384	finger protein Hrf1 -	2.60e+03	467	249	T00169	anti repressor - Stap	2.60e+03
395	4	16.7	153	2	G69426	riboflavin synthase (	2.60e+03	468	251	A55035	cysteine-rich protein	2.60e+03
396	4	16.7	155	2	E11068	hypothetical protein	2.60e+03	469	252	A47188	MHC class II histocom	2.60e+03
397	4	16.7	156	2	E71117	hypothetical protein	2.60e+03	470	254	S50834	emerin - human	2.60e+03
398	4	16.7	159	2	E70379	hypothetical protein	2.60e+03	471	255	S16151	myogenic factor Xmyf-	2.60e+03
399	4	16.7	161	3	T00178	hypothetical protein	2.60e+03	472	255	S22825	Myf-5 protein - mouse	2.60e+03
400	4	16.7	161	3	IS5480	hypothetical natriure	2.60e+03	473	255	JN0624	Myogenic factor - bov	2.60e+03
401	4	16.7	162	2	B71085	hypothetical protein	2.60e+03	474	255	S04105	Myogenic factor 5 - h	2.60e+03
402	4	16.7	162	2	E71239	hypothetical protein	2.60e+03	475	255	A34239	steroid 5alpha-reduct	2.60e+03
403	4	16.7	162	2	S07655	hypothetical protein	2.60e+03	476	256	T00165	repressor - Staphyloc	2.60e+03
404	4	16.7	166	2	A36377	xtwi protein - Africa	2.60e+03	477	258	S41126	Myf5 protein - chicke	2.60e+03
405	4	16.7	167	2	S56670	phosphoprotein phosph	2.60e+03	478	259	S48713	fetal antigen 1 - hum	2.60e+03
406	4	16.7	169	2	J00857	hypothetical 18.1K pr	2.60e+03	479	259	S65744	steroid 5-alpha-reduc	2.60e+03
407	4	16.7	171	2	T00915	probable ACC oxidase	2.60e+03	480	265	A53390	Kunitz-type proteinas	2.60e+03
408	4	16.7	173	2	S72230	transcription factor	2.60e+03	481	267	B33904	CAMP phosphodiesteras	2.60e+03
409	4	16.7	175	2	C70326	hypothetical protein	2.60e+03	482	270	F70432	conserved hypothetical	2.60e+03
410	4	16.7	177	2	A94655	coatamer zeta chain -	2.60e+03	483	272	H71124	hypothetical protein	2.60e+03
411	4	16.7	178	2	E70305	molbdenum cofactor b	2.60e+03	484	273	D70428	hypothetical protein	2.60e+03
412	4	16.7	179	2	S78367	conserved hypothetical	2.60e+03	485	274	C71218	hypothetical protein	2.60e+03
413	4	16.7	180	2	B71070	hypothetical protein	2.60e+03	486	275	A34866	T-cell surface protei	2.60e+03
414	4	16.7	186	2	C71093	hypothetical protein	2.60e+03	487	275	S47325	MyoD protein - zebra	2.60e+03
415	4	16.7	186	2	S08614	cytotoxic T-lymphocyt	2.60e+03	488	275	S08464	T-cell alloantigen RT	2.60e+03
416	4	16.7	186	2	E18783	sodium channel beta 2	2.60e+03	489	275	S51641	MyoD protein - rainbo	2.60e+03
417	4	16.7	189	2	H70416	hypothetical protein	2.60e+03	490	277	I37552	Ox40 homolog - human	2.60e+03
418	4	16.7	190	2	A69382	cytochrome C-type bio	2.60e+03	491	279	A53062	Fas ligand - mouse	2.60e+03
419	4	16.7	195	2	A48190	filamin 1 - human	2.60e+03	492	280	H71320	hypothetical protein	2.60e+03
420	4	16.7	195	2	E39384	finger protein HTF9 -	2.60e+03	493	281	A39484	androgen-withdrawal a	2.60e+03
421	4	16.7	196	2	A39384	finger protein HPF4 -	2.60e+03	494	281	I38707	Fas ligand - human	2.60e+03
422	4	16.7	196	2	A71325	hypothetical protein	2.60e+03	495	283	B71127	probable DNA-3-methyl	2.60e+03
423	4	16.7	197	2	E71025	hypothetical protein	2.60e+03	496	283	T02287	hypothetical protein	2.60e+03
424	4	16.7	200	2	A34632	cysteine dioxygenase	2.60e+03	497	283	B71086	probable 3-methyl-2-o	2.60e+03
425	4	16.7	200	2	S50192	cysteine dioxygenase	2.60e+03	498	287	C34783	myogenesis protein MY	2.60e+03
426	4	16.7	201	2	J66330	twist protein - human	2.60e+03	499	287	S12738	T-cell alloantigen RT	2.60e+03
427	4	16.7	201	2	G01204	twist protein - human	2.60e+03	500	290	C71193	probable dimethyladen	2.60e+03
428	4	16.7	202	2	J66205	cell surface antigen	2.60e+03	501	291	I46539	gamma-sarcoglycan - r	2.60e+03
429	4	16.7	202	2	J64635	tumor-associated L6 a	2.60e+03	502	291	A35002	halorhodopsin - Natro	2.60e+03
430	4	16.7	203	2	A70342	hypothetical protein	2.60e+03	503	292	S33513	gene F1f protein - mo	2.60e+03
431	4	16.7	204	2	I38156	rho protein GDP-disso	2.60e+03	504	293	F70354	ATP-dependent RNA hel	2.60e+03
432	4	16.7	204	2	G02394	fibroblast growth fac	2.60e+03	505	294	F70398	dihydrodipicolinate s	2.60e+03
433	4	16.7	205	2	S55041	multicatalytic endope	2.60e+03	506	296	E71244	hypothetical protein	2.60e+03
434	4	16.7	205	2	C70338	conserved hypothetical	2.60e+03	507	297	A35874	myogenic factor 1 J	2.60e+03
435	4	16.7	205	2	G69379	conserved hypothetical	2.60e+03	508	297	T00635	hypothetical protein	2.60e+03
436	4	16.7	205	2	S40468	proteasome subunit RC	2.60e+03	509	297	T00180	hypothetical protein	2.60e+03
437	4	16.7	206	2	A71019	probable NADH-plastoq	2.60e+03	510	298	A53783	transcription regulat	2.60e+03
438	4	16.7	206	2	C71032	hypothetical protein	2.60e+03	511	298	A32872	myogenic factor CMD1,	2.60e+03
439	4	16.7	206	2	I53066	gene M-twist protein	2.60e+03	512	301	B71330	hypothetical protein	2.60e+03
440	4	16.7	210	2	D70459	hypothetical protein	2.60e+03	513	301	T03320	major capsid protein	2.60e+03
441	4	16.7	212	2	G69369	branched-chain amino	2.60e+03	514	301	D48326	cytb intron 2 protein	2.60e+03
442	4	16.7	215	2	A46245	androgen-induced grow	2.60e+03	515	303	A43708	gamma-interferon-indu	2.60e+03
443	4	16.7	215	2	G20292	fibroblast growth fac	2.60e+03	516	303	A71178	hypothetical protein	2.60e+03
444	4	16.7	215	2	A57843	sodium channel beta 2	2.60e+03	517	303	T00479	hypothetical protein	2.60e+03
445	4	16.7	216	2	J00768	hypothetical 24K prot	2.60e+03	518	305	G70382	phosphatidylglyceroph	2.60e+03
446	4	16.7	217	2	I51062	MHC class II beta cha	2.60e+03	519	306	C70410	hypothetical protein	2.60e+03
447	4	16.7	218	2	E47188	MHC class II histocom	2.60e+03	520	307	S36779	ribosome-binding prot	2.60e+03
448	4	16.7	219	2	A46097	GPI-anchor biosynthes	2.60e+03	521	307	S03716	T-cell receptor beta	2.60e+03
449	4	16.7	223	2	A29063	cytotoxic T-lymphocyt	2.60e+03	522	312	G71131	probable Arp-binding	2.60e+03
450	4	16.7	223	2	I46696	CTLA-4 precursor - ra	2.60e+03	523	312	A56911	TRADD protein - human	2.60e+03
451	4	16.7	224	2	G71281	probable Mg2+ transpo	2.60e+03	524	315	JE0397	RNA polymerase alpha	2.60e+03
452	4	16.7	225	2	A69305	conserved hypothetical	2.60e+03	525	315	C70433	conserved hypothetical	2.60e+03
453	4	16.7	227	2	T00001	pepf protein - Acidia	2.60e+03	526	316	B71301	probable tRNA delta(2	2.60e+03
454	4	16.7	228	2	S51738	translin - human	2.60e+03	527	318	A29636	myoblast determinatio	2.60e+03
455	4	16.7	235	2	E59436	hypothetical protein	2.60e+03	528	318	JC1171	muscle regulatory pro	2.60e+03
456	4	16.7	236	2	E71012	hypothetical protein	2.60e+03	529	319	S26827	myogenic factor 3 - h	2.60e+03
457	4	16.7	238	2	H71291	probable flagellar mo	2.60e+03	530	319	B71312	probable GTP-binding	2.60e+03
458	4	16.7	240	2	JE0011	DNA-directed RNA poly	2.60e+03	531	319	S20086	MyoD1 protein - sheep	2.60e+03
459	4	16.7	244	2	A22722	proliferin-related pr	2.60e+03	532	320	H71259	probable membrane fus	2.60e+03
460	4	16.7	245	2	S63455	hypothetical protein	2.60e+03	533	323	S55348	3',5'-cyclic-nucleoti	2.60e+03
461	4	16.7	245	2	H70438	aminodeoxychorismate	2.60e+03	534	324	T01971	fructokinase (EC 2.7.	2.60e+03



535	4	16.7	325	2	F71283	outer membrane antigen	2.60e+03	608	4	16.7	406	2	A71109	probable phosphate pe	2.60e+03
536	4	16.7	327	2	S56162	MDCR15 protein - huma	2.60e+03	609	4	16.7	406	2	A54857	transcription factor	2.60e+03
537	4	16.7	328	2	S42593	hypothetical protein	2.60e+03	610	4	16.7	406	2	A35484	tryptophan 2,3-dioxyg	2.60e+03
538	4	16.7	328	2	B54058	UDP-N-acetylglucosami	2.60e+03	611	4	16.7	407	2	E71087	hypothetical protein	2.60e+03
539	4	16.7	328	2	F71224	probable 1-aminocyclo	2.60e+03	612	4	16.7	407	2	E71374	probable hemolysin -	2.60e+03
540	4	16.7	330	2	C71180	probable N-acetyl-gam	2.60e+03	613	4	16.7	408	2	F70379	hypothetical protein	2.60e+03
541	4	16.7	334	2	FRHUEP	retrovirus-related pr	2.60e+03	614	4	16.7	410	2	E70475	folylpolyglutamate sy	2.60e+03
542	4	16.7	334	2	F70363	D-lactate dehydrogena	2.60e+03	615	4	16.7	412	2	S43357	inositol 1,4,5-trisph	2.60e+03
543	4	16.7	336	2	F70708	site-specific DNA-met	2.60e+03	616	4	16.7	414	2	S39530	poly(3-hydroxybutyrat	2.60e+03
544	4	16.7	338	2	G71195	hypothetical protein	2.60e+03	617	4	16.7	415	2	F70415	conserved hypothetical	2.60e+03
545	4	16.7	338	2	F71128	hypothetical protein	2.60e+03	618	4	16.7	416	2	B55649	TNFR-associated prote	2.60e+03
546	4	16.7	339	2	B71175	hypothetical protein	2.60e+03	619	4	16.7	417	2	F70391	hypothetical protein	2.60e+03
547	4	16.7	342	2	A40191	platelet-activating f	2.60e+03	620	4	16.7	419	2	E70489	processing proteinase	2.60e+03
548	4	16.7	342	2	H71258	protein tp33 precursor	2.60e+03	621	4	16.7	419	2	A53100	hyaluronate synthase	2.60e+03
549	4	16.7	342	2	S13638	platelet-activating f	2.60e+03	622	4	16.7	419	2	X50187	renin-binding protein	2.60e+03
550	4	16.7	343	2	T01432	secy protein homolog	2.60e+03	623	4	16.7	420	2	S59131	kan-1 protein - rat	2.60e+03
551	4	16.7	344	2	C70401	phospho-N-acetylmuram	2.60e+03	624	4	16.7	420	2	F71040	hypothetical protein	2.60e+03
552	4	16.7	344	2	A69507	conserved hypotheticala	2.60e+03	625	4	16.7	420	2	S55372	DNA-binding protein N	2.60e+03
553	4	16.7	345	2	C55741	thymopoietin gamma pr	2.60e+03	626	4	16.7	422	2	B70444	flagellar hook associ	2.60e+03
554	4	16.7	346	2	A71242	hypothetical protein	2.60e+03	627	4	16.7	422	2	B71147	hypothetical protein	2.60e+03
555	4	16.7	349	2	A41349	histone-specific tran	2.60e+03	628	4	16.7	423	2	A41734	glycine amidinotransf	2.60e+03
556	4	16.7	350	2	A42009	N-formyl peptide rece	2.60e+03	629	4	16.7	423	2	A54140	L-arginine-glycine am	2.60e+03
557	4	16.7	351	2	B42009	FMLP-related receptor	2.60e+03	630	4	16.7	424	2	S15917	developmental control	2.60e+03
558	4	16.7	352	2	A45747	neuropeptide Y/peptid	2.60e+03	631	4	16.7	425	2	B70384	conserved hypotheticala	2.60e+03
559	4	16.7	352	2	G00048	fusin (LESTRA) - crab	2.60e+03	632	4	16.7	425	2	A53798	58K membrane-associat	2.60e+03
560	4	16.7	353	4	S52396	photosystem II protei	2.60e+03	633	4	16.7	428	2	JC4236	V(D)J recombinational	2.60e+03
561	4	16.7	353	2	S28787	neuropeptide Y/peptid	2.60e+03	634	4	16.7	431	2	H70302	conserved hypotheticala	2.60e+03
562	4	16.7	356	2	S55131	hypothetical protein	2.60e+03	635	4	16.7	432	2	A41956	rad9 protein - Schizo	2.60e+03
563	4	16.7	356	2	F71159	hypothetical protein	2.60e+03	636	4	16.7	433	2	S26646	transcription factor	2.60e+03
564	4	16.7	357	2	A39798	RP2 protein, testoste	2.60e+03	637	4	16.7	433	2	S31436	Ig upsilon chain - ax	2.60e+03
565	4	16.7	358	2	G71408	hypothetical protein	2.60e+03	638	4	16.7	436	2	S41782	ABD1 protein - yeast	2.60e+03
566	4	16.7	358	2	S60892	nucleosome assembly p	2.60e+03	639	4	16.7	438	2	G71175	hypothetical protein	2.60e+03
567	4	16.7	359	2	S50532	dpy-20 protein - Caen	2.60e+03	640	4	16.7	438	2	A71026	probable aminotransfe	2.60e+03
568	4	16.7	360	3	JE0362	cytotoxic necrotizing	2.60e+03	641	4	16.7	439	2	S06556	finger protein (clone	2.60e+03
569	4	16.7	361	3	JX0152	D-amino-acid oxidase	2.60e+03	642	4	16.7	439	2	S29319	transcription factor	2.60e+03
570	4	16.7	361	2	A53860	chondroadherin precur	2.60e+03	643	4	16.7	439	2	S29291	transcription factor	2.60e+03
571	4	16.7	363	2	JC4049	polygalacturonase (EC	2.60e+03	644	4	16.7	439	2	E70371	apolipoprotein N-acyl	2.60e+03
572	4	16.7	364	2	F71139	hypothetical protein	2.60e+03	645	4	16.7	441	2	H71064	probable cytoplasmic	2.60e+03
573	4	16.7	364	2	T00717	hypothetical protein	2.60e+03	646	4	16.7	445	2	A34457	202 protein - mouse	2.60e+03
574	4	16.7	366	2	S74005	hypothetical protein	2.60e+03	647	4	16.7	450	2	B71304	probable zinc protein	2.60e+03
575	4	16.7	367	2	F71076	probable hydrogenase	2.60e+03	648	4	16.7	452	2	S55255	lamina-associated pro	2.60e+03
576	4	16.7	368	2	F71204	hypothetical protein	2.60e+03	649	4	16.7	454	2	B55741	thymopoietin beta pre	2.60e+03
577	4	16.7	369	2	A34614	placental protein II	2.60e+03	650	4	16.7	454	2	A46498	glucocorticoid-sensit	2.60e+03
578	4	16.7	370	2	F71221	hypothetical protein	2.60e+03	651	4	16.7	457	2	JC1224	nucleobindin precurs	2.60e+03
579	4	16.7	371	2	H69290	dolichol-P-glucose sy	2.60e+03	652	4	16.7	457	2	A71174	hypothetical protein	2.60e+03
580	4	16.7	371	2	H70326	hypothetical protein	2.60e+03	653	4	16.7	459	2	I55472	Calcium binding prote	2.60e+03
581	4	16.7	372	2	S26667	G protein-coupled rec	2.60e+03	654	4	16.7	461	2	A30222	calcitonin-like recep	2.60e+03
582	4	16.7	374	2	A71320	hypothetical protein	2.60e+03	655	4	16.7	461	2	I60194	intermediate filament	2.60e+03
583	4	16.7	380	2	JQ1679	DNA-binding protein G	2.60e+03	656	4	16.7	464	2	A56600	Na+-dependent phosph	2.60e+03
584	4	16.7	380	2	A71181	probable isomerase -	2.60e+03	657	4	16.7	465	2	I39473	probable 4-aminobuty	2.60e+03
585	4	16.7	383	2	A55739	(MIC) protein MHC cla	2.60e+03	658	4	16.7	466	2	H71234	nucleolar protein Nop	2.60e+03
586	4	16.7	383	2	S53716	homeotic protein dlk	2.60e+03	659	4	16.7	466	2	A55163	probable 4-aminobuty	2.60e+03
587	4	16.7	384	2	F71331	probable cell divisio	2.60e+03	660	4	16.7	467	2	A48916	SOD-2 phosphatase tran	2.60e+03
588	4	16.7	385	2	A54785	preadipocyte factor 1	2.60e+03	661	4	16.7	467	2	S71197	MAD-2 protein homolog	2.60e+03
589	4	16.7	385	2	S54392	major capsid protein	2.60e+03	662	4	16.7	468	2	S71932	myosin VIa - human (	2.60e+03
590	4	16.7	385	2	S53718	homeotic protein dlk	2.60e+03	663	4	16.7	469	2	F70393	hypothetical protein	2.60e+03
591	4	16.7	386	2	S40286	glycine amidinotransf	2.60e+03	664	4	16.7	469	2	C71373	probable oxaloacetate	2.60e+03
592	4	16.7	388	2	C32905	desmoplakin, desmosom	2.60e+03	665	4	16.7	469	2	A57531	EGF alpha transcripti	2.60e+03
593	4	16.7	389	2	F73781	phosphoprotein phosph	2.60e+03	666	4	16.7	469	2	A35789	glutamate-1-semialdeh	2.60e+03
594	4	16.7	390	2	S39780	phosphoprotein phosph	2.60e+03	667	4	16.7	469	2	A55484	p52(Shc) protein - mo	2.60e+03
595	4	16.7	390	2	S65672	phosphoprotein phosph	2.60e+03	668	4	16.7	473	2	S25776	transforming protein	2.60e+03
596	4	16.7	390	2	S20392	phosphoprotein phosph	2.60e+03	669	4	16.7	476	2	JC4646	bone morphogenetic pr	2.60e+03
597	4	16.7	391	2	S54161	L-arginine--glycine a	2.60e+03	670	4	16.7	477	2	B70363	conserved hypotheticala	2.60e+03
598	4	16.7	392	2	PQ0446	DNA-binding protein B	2.60e+03	671	4	16.7	477	2	A42925	hyaluronan receptor R	2.60e+03
599	4	16.7	393	2	I49016	phosphoprotein phosph	2.60e+03	672	4	16.7	478	4	TVHUBD	transforming protein	2.60e+03
600	4	16.7	394	2	G71155	hypothetical protein	2.60e+03	673	4	16.7	478	2	JC1443	hypothetical protein	2.60e+03
601	4	16.7	395	2	A48641	hypothetical protein	2.60e+03	674	4	16.7	478	2	JC4838	bone morphogenetic pr	2.60e+03
602	4	16.7	397	2	JC2524	phosphoprotein phosph	2.60e+03	675	4	16.7	480	2	S52306	zinc finger protein 1	2.60e+03
603	4	16.7	397	2	A71053	hypothetical protein	2.60e+03	676	4	16.7	481	2	F71079	hypothetical protein	2.60e+03
604	4	16.7	398	2	E71454	probable 5-aminolevul	2.60e+03	677	4	16.7	482	2	T01932	RNA binding protein h	2.60e+03
605	4	16.7	401	2	D71003	probable aspartate am	2.60e+03	678	4	16.7	483	2	S41853	centromere/microtubul	2.60e+03
606	4	16.7	402	2	S61413	DNA-binding protein A	2.60e+03	679	4	16.7	483	2	H71162	probable beta-mannos	2.60e+03
607	4	16.7	405	2	S33601	cell adhesion molecul	2.60e+03	680	4	16.7	484	2	S23817	hypothetical protein	2.60e+03

681	4	16.7	490	2	B71338	probable folypolyglu	2.60e+03
682	4	16.7	497	2	G70408	methionine--tRNA liga	2.60e+03
683	4	16.7	498	2	G70429	DNA primase - Aquifex	2.60e+03
684	4	16.7	502	2	G71055	hypothetical protein	2.60e+03
685	4	16.7	505	2	T00400	protein T13E15 6 - Ar	2.60e+03
686	4	16.7	505	2	S62894	alpha-synrophin - hu	2.60e+03
687	4	16.7	505	2	F71218	hypothetical protein	2.60e+03
688	4	16.7	509	2	A48048	egg envelope protein	2.60e+03
689	4	16.7	509	2	T00517	metal ion transporter	2.60e+03
690	4	16.7	510	2	A42750	insulinoma-associated	2.60e+03
691	4	16.7	513	2	A70378	histidine kinase sens	2.60e+03
692	4	16.7	513	2	H71368	hypothetical protein	2.60e+03
693	4	16.7	514	2	A71362	probable nitrogen fix	2.60e+03
694	4	16.7	517	2	A49413	perilipin A - rat	2.60e+03
695	4	16.7	518	2	B48088	beta-transducin repa	2.60e+03
696	4	16.7	518	2	S59642	hypothetical protein	2.60e+03
697	4	16.7	519	2	JC5315	dihydropyrimidinase (	2.60e+03
698	4	16.7	521	2	S57959	LDLP protein - human	2.60e+03
699	4	16.7	521	2	S67491	phosphate transport p	2.60e+03
700	4	16.7	525	2	G71317	hypothetical protein	2.60e+03
701	4	16.7	525	2	H71365	probable licC protein	2.60e+03
702	4	16.7	528	2	A42510	Rag-2 protein - chick	2.60e+03
703	4	16.7	529	2	S52759	zinc metalloproteins	2.60e+03
704	4	16.7	530	2	D70476	DNA helicase - Aquifex	2.60e+03
705	4	16.7	533	2	T00742	ubiquitin-binding pro	2.60e+03
706	4	16.7	534	2	D71344	hypothetical protein	2.60e+03
707	4	16.7	536	2	I67945	phosphodiesterase - r	2.60e+03
708	4	16.7	537	2	I71210	hypothetical protein	2.60e+03
709	4	16.7	537	2	A54424	acrosomal protein Sp3	2.60e+03
710	4	16.7	537	2	S57197	oviduct-specific glyc	2.60e+03
711	4	16.7	538	2	A57624	retinoblastoma protei	2.60e+03
712	4	16.7	538	2	I51560	ribonucleoprotein - A	2.60e+03
713	4	16.7	539	2	I46470	estrogen dependent ov	2.60e+03
714	4	16.7	540	2	F72233	transcription factor	2.60e+03
715	4	16.7	541	2	A48717	glutamate formiminotr	2.60e+03
716	4	16.7	543	2	B54424	acrosomal protein sp3	2.60e+03
717	4	16.7	543	2	S65462	glucose transport pro	2.60e+03
718	4	16.7	544	2	D71243	hypothetical protein	2.60e+03
719	4	16.7	547	2	E69552	signal-transducing hi	2.60e+03
720	4	16.7	547	2	S34744	sterol carrier protei	2.60e+03
721	4	16.7	547	2	F71287	probable treponemal a	2.60e+03
722	4	16.7	548	2	S38864	Ig epsilon chain C re	2.60e+03
723	4	16.7	548	2	A56353	MiD1 protein - yeast	2.60e+03
724	4	16.7	551	2	I45709	endothelial leukocyte	2.60e+03
725	4	16.7	551	2	I46708	endothelial leukocyte	2.60e+03
726	4	16.7	552	2	H70445	hypothetical protein	2.60e+03
727	4	16.7	552	2	I54388	LZFR-1 - human	2.60e+03
728	4	16.7	553	2	A47504	epoxide hydrolase (EC	2.60e+03
729	4	16.7	553	2	T01416	secy protein homolog	2.60e+03
730	4	16.7	554	2	S35587	epoxide hydrolase - h	2.60e+03
731	4	16.7	554	2	A47503	epoxide hydrolase (EC	2.60e+03
732	4	16.7	555	2	JC4711	epoxide hydrolase (EC	2.60e+03
733	4	16.7	556	2	G69327	DNA ligase (lig) homo	2.60e+03
734	4	16.7	560	2	A60164	platelet membrane gly	2.60e+03
735	4	16.7	566	2	C70482	repair excision nucle	2.60e+03
736	4	16.7	567	2	JC5538	Rab geranylgeranyl tr	2.60e+03
737	4	16.7	567	2	A45977	Rab geranylgeranyl tr	2.60e+03
738	4	16.7	569	2	T02435	probable oxysterol-bi	2.60e+03
739	4	16.7	570	2	A57535	intraleukin 1 receptor	2.60e+03
740	4	16.7	572	2	JC5317	dihydropyrimidinase-r	2.60e+03
741	4	16.7	572	2	S49985	TOAB-64 protein - bla	2.60e+03
742	4	16.7	572	2	S58889	collapsin response me	2.60e+03
743	4	16.7	573	2	A33533	cell surface glycopro	2.60e+03
744	4	16.7	574	2	B29677	complement C9 precurs	2.60e+03
745	4	16.7	574	2	A46054	GTP-binding protein A	2.60e+03
746	4	16.7	575	2	JC5432	glycoprotein 6-alpha	2.60e+03
747	4	16.7	576	3	A53982	capsid protein - Euro	2.60e+03
748	4	16.7	580	2	D71229	hypothetical protein	2.60e+03
749	4	16.7	580	2	T02596	hypothetical protein	2.60e+03
750	4	16.7	580	2	A46538	Ig heavy chain, secre	2.60e+03
751	4	16.7	581	2	B54665	netrin-2 precursor--	2.60e+03
752	4	16.7	585	2	A39286	parathyroid hormone /	2.60e+03
753	4	16.7	585	2	I37216	calicrin - human (frag	2.60e+03
754	4	16.7	585	2	S69216	sulfur deprivation re	2.60e+03
755	4	16.7	586	2	T03684	phosphoprotein phosph	2.60e+03
756	4	16.7	588	2	I46003	calicrin - bovine	2.60e+03
757	4	16.7	589	2	A32567	stress-induced protei	2.60e+03
758	4	16.7	592	2	F70455	sulfur oxidation prot	2.60e+03
759	4	16.7	593	2	F70349	NADH dehydrogenase I	2.60e+03
760	4	16.7	595	2	B70402	hypothetical protein	2.60e+03
761	4	16.7	595	2	G02075	transcriptional repress	2.60e+03
762	4	16.7	608	2	T02299	hypothetical protein	2.60e+03
763	4	16.7	610	2	A32558	3',5'-cyclic-nucleoti	2.60e+03
764	4	16.7	610	2	I67946	phosphodiesterase - r	2.60e+03
765	4	16.7	611	2	B71220	hypothetical protein	2.60e+03
766	4	16.7	611	2	F70325	conserved hypothetica	2.60e+03
767	4	16.7	613	2	T00758	ethylene response sen	2.60e+03
768	4	16.7	616	2	I38155	DNA-binding regulator	2.60e+03
769	4	16.7	617	2	F71359	probable polyl-TRNA	2.60e+03
770	4	16.7	618	2	S58513	thyroid iodide transp	2.60e+03
771	4	16.7	621	2	A71221	probable aldehyde:fer	2.60e+03
772	4	16.7	625	2	S13919	potassium channel pro	2.60e+03
773	4	16.7	626	2	A42891	beta-galactosidase (E	2.60e+03
774	4	16.7	629	2	E69284	signal-transducing hi	2.60e+03
775	4	16.7	631	2	JC4298	hyaluronan receptor -	2.60e+03
776	4	16.7	632	2	H70339	NADH dehydrogenase I	2.60e+03
777	4	16.7	633	2	F71146	hypothetical protein	2.60e+03
778	4	16.7	635	2	A54592	110k actin filament-a	2.60e+03
779	4	16.7	637	2	A48189	sodium/phosphate cotr	2.60e+03
780	4	16.7	637	2	A40000	sodium/phosphate cotr	2.60e+03
781	4	16.7	639	2	B48189	sodium/phosphate cotr	2.60e+03
782	4	16.7	641	2	PC4417	Erbb kinase activator	2.60e+03
783	4	16.7	642	2	I46534	renal sodium-dependen	2.60e+03
784	4	16.7	643	2	JC4974	sodium iodide symport	2.60e+03
785	4	16.7	644	2	A42220	helix-loop-helix prot	2.60e+03
786	4	16.7	646	2	H71378	conserved hypothetica	2.60e+03
787	4	16.7	648	2	S16866	gene H19 protein - mo	2.60e+03
788	4	16.7	651	2	I40789	aco operon expression	2.60e+03
789	4	16.7	653	2	S33978	hypothetical protein	2.60e+03
790	4	16.7	654	2	S21552	Mx protein homolog -	2.60e+03
791	4	16.7	656	2	H70365	hypothetical protein	2.60e+03
792	4	16.7	657	2	S32739	filensin - chicken	2.60e+03
793	4	16.7	660	2	T02755	fanconi anemia comple	2.60e+03
794	4	16.7	673	2	S35335	transcription factor	2.60e+03
795	4	16.7	675	3	T00013	DAP-1 beta protein -	2.60e+03
796	4	16.7	679	2	C71007	probable formate dehy	2.60e+03
797	4	16.7	681	2	I78558	hypothetical Brachyur	2.60e+03
798	4	16.7	683	2	I52996	transforming growth f	2.60e+03
799	4	16.7	686	2	A34612	zinc finger protein -	2.60e+03
800	4	16.7	686	2	JC5708	villin-like protein -	2.60e+03
801	4	16.7	686	2	S55788	3',5'-cyclic-nucleoti	2.60e+03
802	4	16.7	686	2	A36317	3',5'-cyclic-nucleoti	2.60e+03
803	4	16.7	692	2	T00025	PSP-95 binding protei	2.60e+03
804	4	16.7	693	2	T00256	hypothetical protein	2.60e+03
805	4	16.7	694	2	G01161	thymopoietin alpha -	2.60e+03
806	4	16.7	694	2	S88442	Grb2-associated binde	2.60e+03
807	4	16.7	694	2	A55741	thymopoietin alpha pr	2.60e+03
808	4	16.7	694	3	T00148	hypothetical protein	2.60e+03
809	4	16.7	702	2	E69498	hypothetical protein	2.60e+03
810	4	16.7	719	2	T00266	hypothetical protein	2.60e+03
811	4	16.7	719	2	S63629	homeotic protein AREC	2.60e+03
812	4	16.7	735	2	A33369	UDPglucose--glycogen	2.60e+03
813	4	16.7	737	2	S18207	adducin alpha splice	2.60e+03
814	4	16.7	737	2	A32156	UDPglucose--glycogen	2.60e+03
815	4	16.7	738	2	S32372	transforming protein	2.60e+03
816	4	16.7	753	2	C71086	hypothetical protein	2.60e+03
817	4	16.7	763	2	A43314	special AT-rich seque	2.60e+03
818	4	16.7	764	2	A56208	DNA-binding protein S	2.60e+03
819	4	16.7	764	2	H70414	conserved hypothetica	2.60e+03
820	4	16.7	767	2	A74639	acetylcholinesterase	2.60e+03
821	4	16.7	772	2	G02869	carnitine O-palmitoyl	2.60e+03
822	4	16.7	772	2	S65532	carnitine palmitoyl	2.60e+03
823	4	16.7	773	3	J50387	sialidase (EC 3.2.1.1	2.60e+03
824	4	16.7	775	2	S63626	homeotic protein AREC	2.60e+03
825	4	16.7	777	2	I48100	ADAM 5 protein precu	2.60e+03
826	4	16.7	778	2	C70412	outer membrane protei	2.60e+03

827	4	16.7	778	2	B71164	probable beta-galacto	2.60e+03	900	4	16.7	1091	2	E71322	probable isoleucyl-tr	2.60e+03
828	4	16.7	785	2	S63552	hypothetical protein	2.60e+03	901	4	16.7	1118	2	S57833	transmembrane protein	2.60e+03
829	4	16.7	786	2	S54773	stat5b protein - mous	2.60e+03	902	4	16.7	1130	2	A48843	MHC class II transact	2.60e+03
830	4	16.7	786	2	I49274	mammary gland factor	2.60e+03	903	4	16.7	1130	2	A56284	differentiation-speci	2.60e+03
831	4	16.7	790	2	S50082	nuclear cap binding p	2.60e+03	904	4	16.7	1131	2	A49393	activator 1 large cha	2.60e+03
832	4	16.7	790	2	A54748	cap-binding protein 8	2.60e+03	905	4	16.7	1131	2	A56200	replication factor C	2.60e+03
833	4	16.7	792	2	T00082	hypothetical protein	2.60e+03	906	4	16.7	1132	2	A35089	myosin-binding protei	2.60e+03
834	4	16.7	793	2	JC5539	Smoothed protein pr	2.60e+03	907	4	16.7	1142	2	S38845	myosin-binding protei	2.60e+03
835	4	16.7	793	2	S54772	mammary gland factor	2.60e+03	908	4	16.7	1142	2	A38968	PI-like adhesin precu	2.60e+03
836	4	16.7	794	2	S55527	mammary gland factor	2.60e+03	909	4	16.7	1146	2	B70376	reverse gyrase - Aqu	2.60e+03
837	4	16.7	794	2	G02317	transcription activat	2.60e+03	910	4	16.7	1147	2	JN0599	DNA-binding protein P	2.60e+03
838	4	16.7	807	2	S40458	ribosomal protein S3	2.60e+03	911	4	16.7	1148	2	A49651	replication factor C	2.60e+03
839	4	16.7	807	2	A34581	oxysterol-binding pro	2.60e+03	912	4	16.7	1148	2	S51855	hypothetical protein	2.60e+03
840	4	16.7	809	2	A34404	oxysterol-binding pro	2.60e+03	913	4	16.7	1152	2	S20106	hypothetical protein	2.60e+03
841	4	16.7	809	2	S40460	ribosomal protein S3	2.60e+03	914	4	16.7	1160	2	T00272	hypothetical protein	2.60e+03
842	4	16.7	816	2	A71006	hypothetical protein	2.60e+03	915	4	16.7	1170	2	A53612	laminin B1 chain pre	2.60e+03
843	4	16.7	823	2	S53647	CDC37 protein - human	2.60e+03	916	4	16.7	1184	2	T00253	gene Ankhn protein -	2.60e+03
844	4	16.7	824	2	I52835	H-NUC - human	2.60e+03	917	4	16.7	1196	2	S35994	DNA repair protein xp	2.60e+03
845	4	16.7	832	2	A40205	Na+/H+-exchanging pro	2.60e+03	918	4	16.7	1196	2	T00715	hypothetical protein	2.60e+03
846	4	16.7	835	2	A40371	DNA-binding protein G	2.60e+03	919	4	16.7	1199	2	A40670	nuclear envelope prot	2.60e+03
847	4	16.7	836	2	JE0248	Arp-binding cassette	2.60e+03	920	4	16.7	1207	2	T00378	KIAA0641 protein - hu	2.60e+03
848	4	16.7	837	2	A42112	mucin-like peptide ML	2.60e+03	921	4	16.7	1211	2	S54500	alpha,alpha-trehalase	2.60e+03
849	4	16.7	837	2	A34898	granulocyte colony-st	2.60e+03	922	4	16.7	1213	2	E69255	mannosyltransferase A	2.60e+03
850	4	16.7	837	2	I57557	DNA-Binding Protein a	2.60e+03	923	4	16.7	1213	2	A41724	limb deformity (ld) p	2.60e+03
851	4	16.7	839	2	A56337	glycoprotein phosphol	2.60e+03	924	4	16.7	1214	2	S28499	probable finger prote	2.60e+03
852	4	16.7	844	2	I53865	phosphodiesterase - i	2.60e+03	925	4	16.7	1215	2	T00364	hypothetical protein	2.60e+03
853	4	16.7	848	2	A54740	interleukin-4-induced	2.60e+03	926	4	16.7	1219	2	I61713	co-repressor protein	2.60e+03
854	4	16.7	850	2	JC5047	ras GTPase-activating	2.60e+03	927	4	16.7	1220	2	A56136	jagged protein precu	2.60e+03
855	4	16.7	851	2	JC5700	ErbB kinase activator	2.60e+03	928	4	16.7	1223	2	I38111	phosphorylase kinase	2.60e+03
856	4	16.7	851	2	A46160	interferon alpha-indu	2.60e+03	929	4	16.7	1228	2	A57384	multimerin, endotheli	2.60e+03
857	4	16.7	853	2	H70470	hypothetical protein	2.60e+03	930	4	16.7	1229	2	A56068	co-repressor protein	2.60e+03
858	4	16.7	855	2	A48168	proliferating-cell nu	2.60e+03	931	4	16.7	1232	2	A55478	neuronal apoptosis in	2.60e+03
859	4	16.7	856	2	G70483	phosphoenolpyruvate s	2.60e+03	932	4	16.7	1233	2	G71612	novel protein kinase	2.60e+03
860	4	16.7	860	2	JC5702	ErbB kinase activator	2.60e+03	933	4	16.7	1234	2	T00363	hypothetical protein	2.60e+03
861	4	16.7	867	2	T00118	ErSH2 protein - sea s	2.60e+03	934	4	16.7	1235	2	S24109	phosphorylase kinase	2.60e+03
862	4	16.7	868	2	JC5701	ErbB kinase activator	2.60e+03	935	4	16.7	1237	2	A31334	phosphorylase kinase	2.60e+03
863	4	16.7	876	2	T00540	probable protein kina	2.60e+03	936	4	16.7	1241	2	S40528	phosphorylase kinase	2.60e+03
864	4	16.7	886	2	A54442	phosphodiesterase - h	2.60e+03	937	4	16.7	1251	2	A56677	neuronal cell cycle w	2.60e+03
865	4	16.7	887	2	A67664	microsomal triglyceri	2.60e+03	938	4	16.7	1268	2	A49674	flightless-I homolog	2.60e+03
866	4	16.7	894	2	I38047	triglyceride transfer	2.60e+03	939	4	16.7	1278	3	A47462	probable DNA-directed	2.60e+03
867	4	16.7	895	2	A55413	triglyceride transfer	2.60e+03	940	4	16.7	1291	2	H71143	probable ribonucleosi	2.60e+03
868	4	16.7	900	2	G71339	probable DNA mismatch	2.60e+03	941	4	16.7	1317	2	T03748	apoptosis associated	2.60e+03
869	4	16.7	901	2	JC5224	methionine--tRNA liga	2.60e+03	942	4	16.7	1352	2	G71051	probable ATP-dependen	2.60e+03
870	4	16.7	902	2	G71286	probable pyruvate, ph	2.60e+03	943	4	16.7	1361	3	T03415	S-layer protein - Cam	2.60e+03
871	4	16.7	907	2	JE0176	orphan G protein-coup	2.60e+03	944	4	16.7	1396	2	A44453	translation initiatio	2.60e+03
872	4	16.7	916	2	I55441	mut15 - rat	2.60e+03	945	4	16.7	1402	2	I46707	initiation factor 4-g	2.60e+03
873	4	16.7	932	2	S47597	mutL protein homolog	2.60e+03	946	4	16.7	1421	2	T00333	hypothetical protein	2.60e+03
874	4	16.7	938	2	A56731	chromatin assembly fa	2.60e+03	947	4	16.7	1427	2	I51669	tumor suppressor - Af	2.60e+03
875	4	16.7	946	3	T00024	ent-kaurene synthase	2.60e+03	948	4	16.7	1429	2	S16233	nitric-oxide synthase	2.60e+03
876	4	16.7	948	2	A57640	retinoblastoma protei	2.60e+03	949	4	16.7	1443	2	I50600	neogenin - chicken (f	2.60e+03
877	4	16.7	951	2	T00017	gene ADAMTS-1 protein	2.60e+03	950	4	16.7	1444	2	S57335	cleavage and polyaden	2.60e+03
878	4	16.7	955	4	C40045	probable transcriptio	2.60e+03	951	4	16.7	1468	2	A44345	nucleoporin - rat	2.60e+03
879	4	16.7	964	2	G71213	probable H(+)-transpo	2.60e+03	952	4	16.7	1475	2	S42718	nuclear pore complex	2.60e+03
880	4	16.7	975	2	I59422	rsec8 - rat (fragment	2.60e+03	953	4	16.7	1493	2	A38218	GAP-associated protei	2.60e+03
881	4	16.7	977	3	T00014	DAP-1 alpha protein -	2.60e+03	954	4	16.7	1510	2	A48227	voltage-dependent Ca2	2.60e+03
882	4	16.7	998	2	T03307	hypothetical protein	2.60e+03	955	4	16.7	1524	2	C71129	pregnancy-associated	2.60e+03
883	4	16.7	991	2	T01372	hypothetical protein	2.60e+03	956	4	16.7	1627	2	S65464	probable reverse gyra	2.60e+03
884	4	16.7	991	2	T00058	hypothetical protein	2.60e+03	957	4	16.7	1646	2	JH0422	voltage-dependent cal	2.60e+03
885	4	16.7	993	2	S35633	DNA-binding protein -	2.60e+03	958	4	16.7	1650	2	S53457	dominant autoantigen	2.60e+03
886	4	16.7	993	2	A71500	Ig mu chain switch re	2.60e+03	959	4	16.7	1651	2	JC1340	outer membrane protei	2.60e+03
887	4	16.7	996	2	A71080	hypothetical protein	2.60e+03	960	4	16.7	1657	2	A54854	Ras GTPase activating	2.60e+03
888	4	16.7	1003	2	T71139	hypothetical protein	2.60e+03	961	4	16.7	1708	2	A05205	hypothetical protein	2.60e+03
889	4	16.7	1017	2	PC4035	cell-cycle-dependent	2.60e+03	962	4	16.7	1719	2	I38902	retinoblastoma-associ	2.60e+03
890	4	16.7	1018	2	JC5799	alpha-D-mannosidase (	2.60e+03	963	4	16.7	1733	2	S27939	tensin - chicken	2.60e+03
891	4	16.7	1027	2	I38759	zinc finger/leucine z	2.60e+03	964	4	16.7	1735	2	A57607	Munc13-1 - rat	2.60e+03
892	4	16.7	1028	2	I37146	myosin I heavy chain	2.60e+03	965	4	16.7	1737	2	T00209	Munc13-1 - rat	2.60e+03
893	4	16.7	1028	2	S41749	myosin heavy chain I	2.60e+03	966	4	16.7	1737	2	A37491	hypothetical helicase	2.60e+03
894	4	16.7	1046	2	A20638	prestalk protein prec	2.60e+03	967	4	16.7	1744	2	A54970	tensin, cardiac muscl	2.60e+03
895	4	16.7	1052	2	T00067	hypothetical protein	2.60e+03	968	4	16.7	1744	2	JH0720	tenascin - African cla	2.60e+03
896	4	16.7	1054	2	D70425	conserved hypothetica	2.60e+03	969	4	16.7	1753	2	T00350	hypothetical protein	2.60e+03
897	4	16.7	1061	2	A57620	steroid receptor coac	2.60e+03	970	4	16.7	1784	2	A49420	tuberos sclerosi pr	2.60e+03
898	4	16.7	1079	2	I59362	calcium/polyvalent ca	2.60e+03	971	4	16.7	1792	2	A57075	tensin - chicken (fa	2.60e+03
899	4	16.7	1085	2	S40476	Ca(2+)-sensing recept	2.60e+03	972	4	16.7	1820	2	A55494	latent transforming g	2.60e+03

```

973 4 16.7 1872 2 T00339 hypothetical protein 2.60e+03
974 4 16.7 1926 2 S01169 beta-glycosidase comp 2.60e+03
975 4 16.7 1938 2 I37453 host cell factor C1 - 2.60e+03
976 4 16.7 1985 2 I61776 Munc13-2 - rat 2.60e+03
977 4 16.7 2035 2 A40718 host cell factor C1 p 2.60e+03
978 4 16.7 2035 2 A56088 host cell factor C1 p 2.60e+03
979 4 16.7 2148 2 A56011 transcription factor 2.60e+03
980 4 16.7 2178 2 S55805 alpha-toxin - Clostri 2.60e+03
981 4 16.7 2549 2 A54837 rapamycin/FKBP12 targ 2.60e+03
982 4 16.7 2549 2 S45340 FKBP-rapamycin-associ 2.60e+03
983 4 16.7 2567 2 A49551 filamin, Mueller cell 2.60e+03
984 4 16.7 2647 2 A37098 gelation factor ABP-2 2.60e+03
985 4 16.7 2649 2 A40937 bullous pemphigoid an 2.60e+03
986 4 16.7 2818 2 B55282 neurofibromatosis-rel 2.60e+03
987 4 16.7 2825 2 I54352 neurofibromin - mouse 2.60e+03
988 4 16.7 2897 2 B48666 cell proliferation an 2.60e+03
989 4 16.7 2825 3 T00133 RNA-directed RNA poly 2.60e+03
990 4 16.7 3085 3 T00327 polyprotein - infecti 2.60e+03
991 4 16.7 3256 2 A48666 cell proliferation an 2.60e+03
992 4 16.7 4096 2 A57099 DNA-activated protein 2.60e+03
993 4 16.7 4128 2 JC6306 protein kinase (EC 2. 2.60e+03
994 4 16.7 4351 2 T00252 MEGF1 protein - rat 2.60e+03
995 4 16.7 4391 2 A38096 perlecan precursor - 2.60e+03
996 4 16.7 4861 2 S71752 giant protein p619 - 2.60e+03
997 4 16.7 4957 2 T03455 ALR protein - human 2.60e+03
998 4 16.7 5262 2 T03454 ALR protein - human 2.60e+03
999 4 16.7 6669 2 S55024 nebulin, skeletal mus 2.60e+03
1000 4 16.7 7962 2 I38346 elastic titin - human 2.60e+03

```

## ALIGNMENTS

```

1 RESULT 1 #type complete
ENTRY #transcription regulator MarR family homolog yfiv - Bacillus
TITLE subtilis
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
24-Sep-1998
H69804 #formal_name Bacillus subtilis
A69880 #formal_name Bacillus subtilis

Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Gollightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Prescan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenecker, T.;

```

```

Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
#journal The complete genome sequence of the Gram-positive bacterium
#title Bacillus subtilis.
#cross-references MUID:98044033
#accession H69804
#status preliminary; nucleic acid sequence not shown;
#molecule_type DNA
#residues 1-160 #label KUN
#cross-references GB:Z99108; GB:AL009126; NID:g2633055; PID:el182831;
#experimental_source strain 168
GENETICS
#gene yfiv
#superfamily transcription regulator yfiv
SUMMARY #length 160 #molecular-weight 18211 #checksum 6771
Query Match 25.0%; Score 6; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 119 GLYKPG 124
QY 12 GLYKPG 17
RESULT 2
ENTRY T02471 #type complete
TITLE hypothetical protein F4118.26 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
#accession T02471
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-279 #label ROU
#cross-references EMBL:AC004665; NID:g3386593; PID:g3386618
GENETICS
#map_position 2
#introns 200/2
#note F4118.26
SUMMARY #length 279 #molecular-weight 31488 #checksum 8843
Query Match 25.0%; Score 6; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 266 GCEPVL 271
QY 17 GCEPVL 22
RESULT 3
ENTRY B70319 #type complete
TITLE conserved hypothetical protein ag_204 - Aquifex aeolicus
ORGANISM #formal_name Aquifex aeolicus
DATE 08-May-1998 #sequence_revision 08-May-1998 #text_change
ACCESSIONS B70319
REFERENCE A70300

```

```

#authors      Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;
              Lenox, A.L.; Graham, D.E.; Overbeek, R.; Sneed, M.A.;
              Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short,
              J.M.; Olson, G.J.; Swanson, R.V.
#journal      Nature (1998) 392:353-358
#title        The complete genome of the hyperthermophilic bacterium
              Aquifex aeolicus.
#cross-references MUID:98196666
#accession     B70319
#status        preliminary; nucleic acid sequence not shown;
              translation not shown
#molecule_type DNA
#residues      1-293 #label AOF
#cross-references GB:AE000678; NID:g2982921; PID:g2982935; GB:AE000657
#experimental_source strain Vfs
GENETICS
#gene          ag_204
SUMMARY
#length 293 #molecular-weight 32952 #checksum 7038
Query Match      25.0%; Score 6; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 174 VTGLYE 179
QY 10 VTGLYE 15
|||||
RESULT 4
ENTRY   I61725 #type complete
TITLE   natural killer associated transcript 2 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE     29-May-1998 #sequence_revision 29-May-1998 #text_change
05-Jun-1998
ACCESSIONS I61725
REFERENCE   A56247
#authors    Colonna, M.; Samaridis, J.
#journal     Science (1995) 268:405-408
#title       Cloning of immunoglobulin-superfamily members associated with
              HLA-C and HLA-B recognition by human natural killer cells.
#cross-references MUID:95232526
#accession   I61725
#status      preliminary; translated from GB/EMBL/DBD
#molecule_type mRNA
#residues    1-341 #label RES
#cross-references GB:L41268; NID:g780305; PID:g780306
GENETICS
#gene        NKAT-2
SUMMARY
#length 341 #molecular-weight 37886 #checksum 1605
Query Match      25.0%; Score 6; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 123 TGLYEX 128
QY 11 TGLYEX 16
|||||
RESULT 5
ENTRY   A47701 #type complete
TITLE   aspartic proteinase ACPL (EC 3.4.23.-) precursor - yeast
              (Candida parapsilosis)
ALTERNATE_NAMES ACP-like acid proteinase ACPL; aspartic proteinase 2
ORGANISM #formal_name Candida parapsilosis
DATE     21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change
13-Sep-1998
ACCESSIONS A47701; S20704; S34581
REFERENCE   A47701
#authors    de Viragh, P.A.; Sanglard, D.; Togni, G.; Falchetto, R.;
              Monod, M.
#journal     J. Gen. Microbiol. (1993) 139:335-342
#title       Cloning and sequencing of two Candida parapsilosis genes
              encoding acid proteases.
#cross-references MUID:93171876
#accession     A47701
#molecule_type DNA
#residues      1-395 #label DE1
#experimental_source isolate CHUV E18
#note          sequence extracted from NCBI backbone (NCBI:125682,
              NCBI:125683)
#note          part of this sequence, including the amino end of the
              mature protein, was confirmed by protein sequencing
REFERENCE     S20704
#authors      Monod, M.
#submission   submitted to the EMBL Data Library, April 1992
#description   Isolation of two tandemly arranged genes for acid protease
              from the yeast Candida parapsilosis.
#accession     S20704
#molecule_type DNA
#residues      1-134, 'R', 135-369, 'A', 372, 'Q', 374-379,
              'LSRSGEVHFCLKCFSLNLETSIVKKAFFNRYIQ', #label MON
#cross-references EMBL:Z11918
#note          This submitted sequence predates the publication A47701,
              differs substantially from it at the carboxyl end, and
              lacks similarity to homologous proteins there
              proteinases.
REFERENCE     S34580
#authors      Fusek, M.; Smith, E.A.; Monod, M.; Foundling, S.I.
#journal       FEBS Lett. (1993) 327:108-112
#title         Candida parapsilosis expresses and secretes two aspartic
              proteinases.
#cross-references MUID:93327903
#accession     S34581
#molecule_type protein
#residues      62-68 #label FUS
#experimental_source isolate CHUV E18
GENETICS
#gene         ACPL
#note          this gene is arranged in tandem with ACPR
#superfamily pepsin
#aspartic proteinase; extracellular protein; glycoprotein;
#hydrolase
FEATURE
1-23      #domain signal sequence #status predicted #label SIG\
24-61     #domain propeptide #status predicted #label PRO\
62-395    #product aspartic proteinase 2 #status experimental
              #label MAT\
93,272    #active_site Asp #status predicted
SUMMARY
#length 395 #molecular-weight 42050 #checksum 2709
Query Match      25.0%; Score 6; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 334 EFVLQS 339
QY 19 EFVLQS 24
|||||
RESULT 6
ENTRY   I61727 #type complete
TITLE   natural killer associated transcript 4 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE     29-May-1998 #sequence_revision 29-May-1998 #text_change
05-Jun-1998
ACCESSIONS I61727
REFERENCE   A56247
#authors    Colonna, M.; Samaridis, J.
#journal     Science (1995) 268:405-408
#title       Cloning of immunoglobulin-superfamily members associated with
              HLA-C and HLA-B recognition by human natural killer cells.
#cross-references MUID:95232526
#accession   I61727
#status      preliminary; translated from GB/EMBL/DBD
#molecule_type mRNA
#residues    1-455 #label RES

```

```
##cross-references GB:L41270; NID:g780309; PID:g780310
GENETICS
#gene NKAT-4
#length 455 #molecular-weight 50230 #checksum 5605
SUMMARY
Query Match 25.0%; Score 6; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 218 TGLYEK 223
Qy 11 TGLYEK 16
RESULT 7
ENTRY #type complete
TITLE KIR (CL-5) NK receptor protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change
17-Jul-1998
ACCESSIONS G01923
REFERENCE G08778
#authors Wagtman, N.
#submission submitted to the EMBL Data Library, June 1995
#accession G01923
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-455 #label WAG
#cross-references EMBL:U30272; NID:g1004356; PID:g1004357
SUMMARY #length 455 #molecular-weight 50216 #checksum 5561
Query Match 25.0%; Score 6; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 218 TGLYEK 223
Qy 11 TGLYEK 16
RESULT 8
ENTRY #type complete
TITLE tRNA-pseudouridine synthase I (EC 5.4.99.12) PUS1 - yeast
ALTERNATE_NAMES (Saccharomyces cerevisiae)
ORGANISM protein p1805; protein YPL212c
#formal_name Saccharomyces cerevisiae
DATE 10-Dec-1994 #sequence_revision 31-May-1996 #text_change
06-Feb-1998
ACCESSIONS S65231; S69218; S69220
REFERENCE S65202
#authors Rieger, M.; Mueller-Auer, S.; Schaefer, M.
#submission submitted to the Protein Sequence Database, May 1996
#accession S65231
#molecule_type DNA
#residues 1-544 #label RIE
#cross-references EMBL:U73568; NID:g1370439; PID:e247066; PID:g1370440;
MIPS:YPL212c
#experimental_source strain S288C (AB972)
REFERENCE S69218
#authors Hurt, E.C.
#submission submitted to the EMBL Data Library, July 1994
#accession S69218
#molecule_type DNA
#residues 1-544 #label HUR
#cross-references EMBL:X80673; NID:g1403090; PID:e113803; PID:g1403091
S69220
#authors Simos, G.; Tekotte, H.; Grosjean, H.; Segref, A.; Sharma, K.;
Tollervey, D.; Hurt, E.C.
#journal EMBO J. (1996) 15:2270-2284
#title Nuclear pore proteins are involved in the biogenesis of
functional tRNA.
#cross-references MUID:96208513
#accession S69220
##status nucleic acid sequence not shown
##molecule_type DNA
##residues 77-205; 306-446 #label SIM
##cross-references EMBL:X80673
GENETICS
#gene SGD:PUS1
#cross-references SGD:S0006133; MIPS:YPL212c
#map_position 16L
KEYWORDS intramolecular transferase; isomerase; nucleus
SUMMARY #length 544 #molecular-weight 62142 #checksum 6840
Query Match 25.0%; Score 6; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 463 FNKVTG 468
Qy 7 FNKVTG 12
RESULT 9
ENTRY #type complete
TITLE hypothetical protein YHR109w - yeast (Saccharomyces
cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change
12-Dec-1997
ACCESSIONS S48951
REFERENCE S46691
#authors Latreille, P.
#submission submitted to the EMBL Data Library, May 1994
#description The sequence of S. cerevisiae cosmid 8263.
#accession S48951
#molecule_type DNA
#residues 1-585 #label LAT
#cross-references EMBL:U00059; NID:g529116; PID:g529122; MIPS:YHR109w
GENETICS
#map_position 8R
SUMMARY #length 585 #molecular-weight 68284 #checksum 7679
Query Match 25.0%; Score 6; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 60 KRCTFN 65
Qy 3 KRCTFN 8
RESULT 10
ENTRY #type complete
TITLE 165K myofibrillar M-band structural protein - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change
13-Sep-1998
ACCESSIONS A44027
REFERENCE A44027
#authors Noguchi, J.; Yanagisawa, M.; Imamura, M.; Kasuya, Y.;
Sakurai, T.; Tanaka, T.; Masaki, T.
#journal J. Biol. Chem. (1992) 267:20302-20310
#title Complete primary structure and tissue expression of chicken
pectoralis M-protein.
#cross-references MUID:93015907
#accession A44027
#status preliminary
#molecule_type mRNA; protein
#residues 1-1450 #label NOG
#cross-references GB:D11474; NID:g222832; PID:d1002515; PID:g222833
#experimental_source pectoralis muscle
#note sequence extracted from NCBI backbone (NCBIN:115719,
NCBIP:115720)
CLASSIFICATION #superfamily skelemin
SUMMARY #length 1450 #molecular-weight 163400 #checksum 4213
```

Query Match 25.0%; Score 6; DB 2; Length 1450;  
Best Local Similarity 100.0%; Pred. No. 2.60e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 432 VTGLYE 437  
QY 10 VTGLYE 15  
|||||

RESULT 11  
ENTRY I38344 #type complete  
TITLE titin, cardiac muscle - human  
ALTERNATE\_NAMES connectin  
CONTAINS serine/threonine-specific protein kinase (EC 2.7.1.-)  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 02-Jul-1998

ACCESSIONS I38344; I38345; S20898; S20897; S20899; S63665; S37393  
REFERENCE A57430  
#authors Labeit, S.; Kolmerer, B.  
#journal Science (1995) 270:293-296  
#title Titins: giant proteins in charge of muscle ultrastructure and elasticity.  
#cross-references MUID:96026330  
#accession I38344  
#status nucleic acid sequence not shown; translation not shown;  
translated from GB/EMBL/DBJ

#molecule\_type mRNA  
#residues 1-26926 #label LAB1  
#cross-references EMBL:X90568; NID:g1017424; PID:g1017425  
REFERENCE I38345  
#authors Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.  
#journal Biochemistry (1995) 34:553-561  
#title Bisecting titin into its structural motifs: identification of an alpha-helix motif near the titin N-terminus.  
#cross-references MUID:95119041  
#accession I38345  
#status nucleic acid sequence not shown

#molecule\_type mRNA  
#residues 1977-2014 #label MUS  
#cross-references EMBL:X83270; NID:g602579; PID:g602580  
#note conformation and properties are reported for a synthetic peptide corresponding to the translated fragment shown

REFERENCE S20897  
#authors Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.  
#journal EMBO J. (1992) 11:1711-1716  
#title Towards a molecular understanding of titin.  
#cross-references MUID:92258380  
#accession S20898  
#status nucleic acid sequence not shown

#molecule\_type mRNA  
#residues 13597-14200,'I',14202-14696 #label LAB2  
#cross-references EMBL:X64698; NID:g37192; PID:g37193  
#accession S20897  
#status nucleic acid sequence not shown; translation not shown

#molecule\_type mRNA  
#residues 16330-16382,'S',16384-16756,'F',16758-16860 #label LAB3  
#cross-references EMBL:X64699; NID:g37190; PID:g37191  
#accession S20899  
#status nucleic acid sequence not shown; translation not shown

#molecule\_type mRNA  
#residues 'P',22278-22431,'R',22433-22448,'G',22450-22453,'O',22455-22480,'TR',22483-22669,'N',22671-22696,'SA',22699-23323,'L',23325-25376 #label LAB4  
#cross-references EMBL:X64697; NID:g37190; PID:g37195  
REFERENCE S63665  
#authors Kolmerer, B.; Olivieri, N.; Witt, C.C.; Herrmann, B.G.; Labeit, S.  
#journal J. Mol. Biol. (1996) 256:556-563  
#title Genomic organization of M line titin and its tissue-specific expression in two distinct isoforms.  
#cross-references MUID:96177761

#accession S63665  
#status nucleic acid sequence not shown  
#molecule\_type DNA  
#residues 26729-26825 #label KOL  
#cross-references EMBL:X92412; NID:g1236761  
REFERENCE S37393  
#authors Gautel, M.; Leonard, K.; Labeit, S.  
#journal EMBO J. (1993) 12:3827-3834  
#title Phosphorylation of KSP motifs in the C-terminal region of titin in differentiating myoblasts.  
#cross-references MUID:94008990  
#accession S37393  
#molecule\_type mRNA  
#residues 26831-26926 #label GAU  
REFERENCE A66736  
#authors Impta, S.; Politou, A.S.; Pastore, A.  
#submission submitted to the Brookhaven Protein Data Bank, February 1996  
#cross-references PDB:1TIT  
#contents annotation; conformation by (1)H-NMR, residues 5253-5341  
REFERENCE A66201  
#authors Pfuhl, M.; Pastore, A.  
#submission submitted to the Brookhaven Protein Data Bank, August 1996  
#cross-references PDB:1NCT  
#contents annotation; conformation by (1)H-NMR, residues 'S', 26059-26155  
GENETICS  
#gene GDB:TTN  
#cross-references GDB:127867; OMIM:188840  
#map\_position q31-2q32  
FUNCTION structural protein forming filaments in striated muscle  
#superfamily\_titin; fibronectin type III repeat homology; immunoglobulin homology; protein kinase homology  
CLASSIFICATION  
KEYWORDS alternative splicing; calmodulin binding; cardiac muscle; duplication; glycoprotein; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase; skeletal muscle; structural protein  
FEATURE  
24752-25008 #domain protein kinase homology #label KIN  
84,177,905,2276,  
2378,2459,2481,  
2563,2669,2763,  
2896,3088,3179,  
3384,3432,3628,  
3772,4068,4318,  
5047,5246,5823,  
6213,6264,6699,  
6800,7083,7300,  
7506,7597,8338,  
8447,8455,8719,  
8938,9375,10130,  
10133,10177,10398,  
11066,11488,11515,  
11635,11949,12170,  
12478,12526,12845,  
12875,13001,13036,  
13295,13540,13787,  
13986,14085,14529,  
14870,14910,15438,  
15704,16397,16520,  
16695  
#binding\_site carbohydrate (Asn) (covalent) #status predicted\

```
21900,21935,22295,
22495,22627,22897,
23024,23318,23883,
24012,24177,24290,
24447,24642,24840,
25201,25927,26133,
26283,26682,26698,
26811
#binding_site carbohydrate (Asn) (covalent) #status
predicted\
26171,26178,26184,
#binding_site phosphate (Ser) (covalent) #status
experimental
SUMMARY #length 26926 #molecular-weight 2993498 #checksum 431
Query Match 25.0%; Score 6; DB 1; Length 26926;
Best Local Similarity 100.0%; Pred. No. 2.60e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 11267 VTGLYE 11272
|||||
Qy 10 VTGLYE 15

RESULT 12
ENTRY #type complete
TITLE transferrin-binding protein - Trypanosoma brucei
ORGANISM #formal_name Trypanosoma brucei
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
24-Jul-1997
ACCESSIONS S15030
REFERENCE Schell, D.; Evers, R.; Preis, D.; Ziegelbauer, K.; Kiefer,
H.; Lottspeich, F.; Cornelissen, A.W.C.A.; Overath, P.
EMBO J. (1991) 10:1061-1066
#journal A transferrin-binding protein of Trypanosoma brucei is
#title encoded by one of the genes in the variant surface
glycoprotein gene expression site.
#cross-references MUID:91216099
#accession S15030
#status preliminary
#molecule_type protein
#residues 1-27 #label SCH
SUMMARY #length 27 #molecular-weight 2888 #checksum 8838
Query Match 20.8%; Score 5; DB 2; Length 27;
Best Local Similarity 83.3%; Pred. No. 1.08e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 14 NKVXGL 19
|||||
Qy 8 NKVTGL 13

RESULT 13
ENTRY #type complete
TITLE hypothetical protein BB0424 - Lyme disease spirochete
ORGANISM #formal_name Borrelia burgdorferi #common_name Lyme disease
spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
05-Jun-1998
ACCESSIONS G70152
REFERENCE J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Utterback, T.; Watthey, L.; McDonald, L.; Artiach, P.;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
Nature (1997) 390:580-586
#journal Genomic sequence of a Lyme disease spirochaete, Borrelia
```

```
burgdorferi.
#cross-references MUID:98065943
#accession G70152
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-49 #label KLE
#cross-references GB:AE001147; GB:AE000783; NID:g2688330; PID:g2688334;
TIGR:BB0424
SUMMARY #length 49 #molecular-weight 5602 #checksum 5497
Query Match 20.8%; Score 5; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.08e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 45 FNKYT 49
|||||
Qy 7 FNKYT 11

RESULT 14
ENTRY #type complete
TITLE monellin chain B - serendipity berry
ALTERNATE_NAMES monellin chain II
ORGANISM #formal_name Dioscoreophyllum cumminsii #common_name
serendipity berry
DATE 24-Apr-1984 #sequence_revision 25-Oct-1996 #text_change
08-May-1998
ACCESSIONS JH0210; B91670; B90607; A03376
REFERENCE JH0209
#authors Kohmura, M.; Nio, N.; Ariyoshi, Y.
#journal Agric. Biol. Chem. (1990) 54:2219-2224
#title Complete amino acid sequence of the sweet protein monellin.
#accession JH0210
#molecule_type protein
#residues 1-51 #label KOH
#cross-references MUID:77004858
#accession B91670
#molecule_type protein
#residues 2-49, 'NE' #label FRA
REFERENCE A91670
#authors Frank, G.; Zuber, H.
#journal Hoppe-Seyler's Z. Physiol. Chem. (1976) 357:585-592
#title The complete amino acid sequences of both subunits of the
sweet protein monellin.
#cross-references MUID:77004858
#accession B91670
#molecule_type protein
#residues 2-49, 'NE' #label FRA
REFERENCE A90607
#authors Bohak, Z.; Li, S.L.
#journal Biochim. Biophys. Acta (1976) 427:153-170
#title The structure of monellin and its relation to the sweetness
of the protein.
#cross-references MUID:76161292
#accession B90607
#molecule_type protein
#residues 2-51 #label BOH
#note this sequence is called chain II by the authors
REFERENCE A51514
#authors Jiang, F.; Tong, L.; Kim, S.H.
#submission submitted to the Brookhaven Protein Data Bank, August 1992
#cross-references PDB:3MON
#contents annotation: X-ray crystallography, 2.8 angstroms, residues
2-49, 'NE'
REFERENCE A58205
#authors Ogata, C.; Hatada, M.; Tomlinson, G.; Shin, W.C.; Kim, S.H.
#journal Nature (1987) 328:739-742
#title Crystal structure of the intensely sweet protein monellin.
#cross-references MUID:87287292
#contents annotation: X-ray crystallography, 3.0 angstroms, residues
2-49, 'NE'
COMMENT It is possible that both chains are derived from a single precursor
```



```

chain.
COMMENT  Blocking of the active site cysteine abolishes the sweet taste.
COMPLEX  heterodimer (see PIR:MLDIA)
CLASSIFICATION
KEYWORDS  #superfamily monellin chain B
          heterodimer; sweet taste
FEATURE
1-51      #product monellin chain Thr-B #status experimental
          #label MAT3\
2-51      #product monellin chain B #status experimental #label
          MAT1\
3-51      #product monellin chain des-Gly-B #status experimental
          #label MAT2\
42         #active_site Cys #status experimental
SUMMARY   #length 51 #molecular-weight 5956 #checksum 1001
Query Match      20.8%; Score 5; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.08e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 TFNKV 38
    |||||
QY 6 TFNKV 10

RESULT 15
ENTRY
TITLE      NTKN6G      #type complete
           omega-conotoxin GVIB precursor - cone shell (Conus
           geographus)
ALTERNATE_NAMES  shaker peptide GVIB
CONTAINS         omega-conotoxin GVIA; omega-conotoxin GVIC
ORGANISM         #formal_name Conus geographus #common_name geography cone
DATE             23-Feb-1985 #sequence_revision 23-Mar-1995 #text_change
                23-May-1997
ACCESSIONS      A44006; A60133; B60133; A01785
REFERENCE
A44006         Colledge, C.J.; Hunsperger, J.P.; Imperial, J.S.; Hillyard,
#authors       D.R.
#journal       Toxicol (1992) 30:1111-1116
#title         Precursor structure of omega-conotoxin GVIA determined from a
               CDNA clone.
#cross-references MUID:93069266
#accession     A44006
               #molecule_type mRNA
               ##residues 1-73 #label COL
               ##cross-references GB:M84612; NID:g156520; PID:g1070393
               ##experimental_source venom duct
               ##note      sequence extracted from NCBI backbone (NCBIN:119531,
               NCBIP:119532)
REFERENCE      A43620
#authors       Olivera, B.M.; Gray, W.R.; Zeikus, R.; McIntosh, J.M.; Varga,
               J.; Rivier, J.; de Santos, V.; Cruz, L.J.
#journal       Science (1985) 230:1338-1343
#title         Peptide neurotoxins from fish-hunting cone snails.
#cross-references MUID:86070213
#accession     A60133
               ##molecule_type protein
               ##residues 46-73 #label OLI
               #accession B60133
               ##molecule_type protein
               ##residues 46-71 #label OLI
REFERENCE      A01785
#authors       Olivera, B.M.; McIntosh, J.M.; Cruz, L.J.; Luque, F.A.; Gray,
               W.R.
#journal       Biochemistry (1984) 23:5087-5090
#title         Purification and sequence of a presynaptic peptide toxin from
               Conus geographus venom.
#cross-references MUID:85072796
#accession     A01785
               ##molecule_type protein
               ##residues 46-72 #label OL3
REFERENCE      A49017
#authors       Nishiuchi, Y.; Kumagaye, K.; Noda, Y.; Watanabe, T.X.;
               Sakakibara, S.

```

```

#journal      Biopolymers (1986) 25:S61-S68
#title        Synthesis and secondary-structure determination of
               omega-conotoxin GVIA: a 27-peptide with three
               intramolecular disulfide bonds.
#contents     annotation
#note         disulfide bonds determined and confirmed by chemical
               synthesis
REFERENCE      A51894
#authors      Davis, J.H.; Bradley, E.K.; Miljanich, G.P.; Nadasdi, L.;
               Ramachandran, J.; Basus, V.J.
#submission   submitted to the Brookhaven Protein Data Bank, April 1993
#cross-references PDB:1OMC
#contents     annotation; conformation by (1)H-NMR, residues 46-72
REFERENCE      A58536
#authors      Davis, J.H.; Bradley, E.K.; Miljanich, G.P.; Nadasdi, L.;
               Ramachandran, J.; Basus, V.J.
#journal      Biochemistry (1993) 32:7396-7405
#title        Solution structure of omega-conotoxin GVIA using 2-D NMR
               spectroscopy and relaxation matrix analysis.
#contents     annotation; conformation by (1)H-NMR
REFERENCE      A51089
#authors      Pallaghy, P.K.; Duggan, B.M.; Pennington, M.W.; Norton, R.S.
#submission   submitted to the Brookhaven Protein Data Bank, August 1993
#cross-references PDB:1CCO
#contents     annotation; conformation by (1)H-NMR, residues 46-72
COMMENT       There are several types of conotoxins: alpha, acting on
               postsynaptic membranes; omega, acting at presynaptic membranes;
               and mu, acting on muscle membranes. Conotoxin GVIA is an
               omega-type neurotoxin.
CLASSIFICATION #superfamily omega-conotoxin
KEYWORDS       acetylcholine release inhibition; amidated carboxyl end;
               calcium channel inhibitor; hydroxyproline; presynaptic
               neurotoxin; venom
FEATURE
1-22         #domain signal sequence #status predicted #label SIG\
23-45         #domain propeptide #status predicted #label PRO\
46-73         #product omega-conotoxin GVIB #status experimental
               #label MAT1\
46-72         #product omega-conotoxin GVIA #status experimental
               #label MAT2\
46-71         #product omega-conotoxin GVIC #status experimental
               #label MAT3\
46-61,53-64,60-71 #disulfide_bonds #status experimental\
49,55,66       #modified_site 4-hydroxyproline (pro) #status
               experimental\
72            #modified_site amidated carboxyl end (Tyr) (amide in
               mature form from following glycine) #link MAT2 #status
               experimental
SUMMARY        #length 73 #molecular-weight 7851 #checksum 9169
Query Match    20.8%; Score 5; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.08e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 YTKRC 71
    |||||
QY 1 YTKRC 5

Search completed: Sat Aug 28 14:44:40 1999
Job time : 43 secs.

```

**This Page Blank (uspto)**

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Result No.	Query		Length	DB	ID	Description	Pred. No.	
	Score	Match						
1	7	29.2	387	1	NK10_HUMAN	MHC CLASS I NK CELL RE	1.41e-02	
2	6	25.0	211	1	U192_HSV7J	PROTEIN U63.	1.16e+00	
3	6	25.0	301	1	NKR5_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
4	4	25.0	304	1	NKR7_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
5	6	25.0	304	1	NKR9_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
6	6	25.0	341	1	NKR2_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
7	6	25.0	348	1	NKR0_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
8	6	25.0	412	1	CAR2_CANPA	CANDIDAPEPSIN 2 PRECUR	1.16e+00	
9	6	25.0	455	1	NKR4_HUMAN	MHC CLASS I NK CELL RE	1.16e+00	
10	6	25.0	544	1	PUS1_YEAST	PSEUDOURIDYLATE SYNTHA	1.16e+00	
11	6	25.0	585	1	YHQ9_YEAST	HYPOTHEICAL 68.3 KD P	1.16e+00	
12	6	25.0	903	1	STLM_HUMAN	PROBABLE LEUCYL-TRNA S	1.16e+00	
13	6	25.0	1450	1	MP5F_CHICK	M-PROTEIN, STRIATED MU	1.16e+00	
14	5	20.8	43	1	RL11_STRGB	50S RIBOSOMAL PROTEIN	6.13e+01	
15	5	20.8	5	1	MONB_DIOCU	MONELLIN, CHAIN B (CHA	6.13e+01	
16	5	20.8	61	1	PD1L_BOVIN	PHOSPHODIESTERASE I (E	6.13e+01	
17	5	20.8	91	1	XKP6_CONGE	OMEGA-CONOTOXINS GVIA,	6.13e+01	
18	5	20.8	73	1	XP07_BACLI	XPART7 PROTEIN	6.13e+01	
19	5	20.8	123	1	Y342_METUA	HYPOTHEICAL PROTEIN M	6.13e+01	
20	5	20.8	129	1	G513_BACSV	GENERAL STRESS PROTEIN	6.13e+01	
21	5	20.8	132	1	RS24_XENIA	40S RIBOSOMAL PROTEIN	6.13e+01	
22	5	20.8	133	1	RS24_HUMAN	40S RIBOSOMAL PROTEIN	6.13e+01	
23	5	20.8	139	1	YM28_MARPO	HYPOTHEICAL 16.2 KD P	6.13e+01	

97	5	20.8	337	1	YKOT_BACSU	HYPOTHETICAL 38.5 KD P	6.13e+01	170	5	20.8	647	1	COAT_ADVG	COAT PROTEIN VP1 [CONT	6.13e+01
98	5	20.8	338	1	YP79_CAEEL	HYPOTHETICAL 40.2 KD P	6.13e+01	171	5	20.8	651	1	SYM_METH	METHIONYL-TRNA SYNTHET	6.13e+01
99	5	20.8	341	1	EPG_STPRA	ELONGATION FACTOR G (E	6.13e+01	172	5	20.8	654	1	PSTA_MYCGE	PHOSPHATE TRANSPORT SY	6.13e+01
100	5	20.8	348	1	NK1_HUMAN	MHC CLASS I NK CELL RE	6.13e+01	173	5	20.8	658	1	CPT2_HUMAN	MITOCHONDRIAL CARNITIN	6.13e+01
101	5	20.8	350	1	R51B_MOUSE	DNA REPAIR PROTEIN RAD	6.13e+01	174	5	20.8	689	1	PYS2_PSEAE	PYOCIN S2 (EC 3.1.21.1	6.13e+01
102	5	20.8	350	1	R51B_HUMAN	DNA REPAIR PROTEIN RAD	6.13e+01	175	5	20.8	701	1	EPG_MYCLE	ELONGATION FACTOR G (E	6.13e+01
103	5	20.8	359	1	OCD_BRUBA	ORNITHINE CYCLODEAMINA	6.13e+01	176	5	20.8	721	1	ERG7_SCHPO	LANOSTEROL SYNTHASE (E	6.13e+01
104	5	20.8	361	1	GLR_YEAST	POLYGALACTURONASE PREC	6.13e+01	177	5	20.8	725	1	MYT1_HUMAN	MYELIN TRANSCRIPTION F	6.13e+01
105	5	20.8	393	1	BCA1_YEAST	BRANCHED-CHAIN AMINO A	6.13e+01	178	5	20.8	745	1	SVL_CAEEL	PROBABLE GLYCYL-TRNA S	6.13e+01
106	5	20.8	399	1	P2X1_HUMAN	P2X PURINOCEPTOR 1 (AT	6.13e+01	179	5	20.8	762	1	PMIP_SCHPO	PROBABLE MITOCHONDRIAL	6.13e+01
107	5	20.8	400	1	LMB_HIRME	LAMININ B-CHAIN (FRAGM	6.13e+01	180	5	20.8	764	1	SYFB_HELPY	PHENYLALANYL-TRNA SYN	6.13e+01
108	5	20.8	404	1	Y4XM_RHIS	HYPOTHETICAL TRANSPORT	6.13e+01	181	5	20.8	776	1	VP4_ROT6	OUTER CAPSID PROTEIN V	6.13e+01
109	5	20.8	405	1	CAR8_CANAL	CANDIDAPEPSIN 8 PRECUR	6.13e+01	182	5	20.8	783	1	NQO3_THETH	NADH-UBIQUINONE OXIDOR	6.13e+01
110	5	20.8	409	1	ODPA_SCHPO	PYRUVATE DEHYDROGENASE	6.13e+01	183	5	20.8	788	1	DEOL_HPBEH	DNA POLYMERASE (EC 2.7	6.13e+01
111	5	20.8	410	1	Y457_METJA	HYPOTHETICAL PROTEIN M	6.13e+01	184	5	20.8	804	1	SVL_BACSU	LEUCYL-TRNA SYNTHETASE	6.13e+01
112	5	20.8	413	1	HMO_WYACE	HOMOLIN PRECURSOR (P4	6.13e+01	185	5	20.8	810	1	YLZ5_CAEEL	HYPOTHETICAL 89.2 KD P	6.13e+01
113	5	20.8	414	1	GLTA_CAWJE	SERINE HYDROXYMETHYLTR	6.13e+01	186	5	20.8	853	1	ACZ5_CAEEL	ACONITATE HYDRATASE 2	6.13e+01
114	5	20.8	418	1	PGK_EUPCR	PHOSPHOGLYCERATE KINAS	6.13e+01	187	5	20.8	859	1	CHS1_SCHPO	CHITIN SYNTHASE 1 (EC	6.13e+01
115	5	20.8	419	1	PGK_OXINO	PHOSPHOGLYCERATE KINAS	6.13e+01	188	5	20.8	871	1	PC1_MOUSE	PLASMA-CELL MEMBRANE G	6.13e+01
116	5	20.8	419	1	DHE3_THELI	GLUTAMATE DEHYDROGENAS	6.13e+01	189	5	20.8	873	1	PC1_HUMAN	PLASMA-CELL MEMBRANE G	6.13e+01
117	5	20.8	420	1	ILSR_HUMAN	INTERLEUKIN-5 RECEPTOR	6.13e+01	190	5	20.8	878	1	SVL_TREPA	LEUCYL-TRNA SYNTHETASE	6.13e+01
118	5	20.8	422	1	LACE_AGRDD	LACTOSE-BINDING PROTEI	6.13e+01	191	5	20.8	880	1	SVY_BACSU	VALYL-TRNA SYNTHETASE	6.13e+01
119	5	20.8	422	1	NCAP_CHAV	NUCLEOCAPSID PROTEIN (	6.13e+01	192	5	20.8	903	1	ECIC_BOVIN	EPITHELIAL CHLORIDE CH	6.13e+01
120	5	20.8	427	1	GLYA_SYNY3	SERINE HYDROXYMETHYLTR	6.13e+01	193	5	20.8	953	1	Y635_HAEIN	PROBABLE TONB-DEPENDEN	6.13e+01
121	5	20.8	429	1	YDIS_ECOLI	PROBABLE ELECTRON TRAN	6.13e+01	194	5	20.8	964	1	UL70_MCMVS	HELICASE/PRIMASE COMPL	6.13e+01
122	5	20.8	433	1	Y384_MYCGE	PROBABLE GTP-BINDING P	6.13e+01	195	5	20.8	986	1	UGNZ_CLOSR	ENDOGLUCANASE Z PRECUR	6.13e+01
123	5	20.8	435	1	TEF5_HUMAN	TRANSCRIPTIONAL ENHANC	6.13e+01	196	5	20.8	1017	1	LRG1_YEAST	LRG1 PROTEIN	6.13e+01
124	5	20.8	439	1	TEF5_MOUSE	TRANSCRIPTIONAL ENHANC	6.13e+01	197	5	20.8	1035	1	NRAN_STRPN	SIALIDASE A PRECURSOR	6.13e+01
125	5	20.8	440	1	SCAL_DROME	SCALLOPED PROTEIN.	6.13e+01	198	5	20.8	1039	1	ITAB_HUMAN	PLATELET MEMBRANE GLYC	6.13e+01
126	5	20.8	445	1	SH7_HUMAN	5-HYDROXYTRYPTAMINE 7	6.13e+01	199	5	20.8	1053	1	ITG5_MOUSE	HYPOTHETICAL RECEPTOR	6.13e+01
127	5	20.8	446	1	BRNO_LACDL	BRANCHED-CHAIN AMINO A	6.13e+01	200	5	20.8	1123	1	VG39_HSVI1	HYPOTHETICAL GENE 39 P	6.13e+01
128	5	20.8	448	1	SH7_RAT	5-HYDROXYTRYPTAMINE 7	6.13e+01	201	5	20.8	1128	1	PHYA_ORISA	PHYTOCHROME A.	6.13e+01
129	5	20.8	448	1	SH7_MOUSE	5-HYDROXYTRYPTAMINE 7	6.13e+01	202	5	20.8	1131	1	PHYA_MAIZE	PHYTOCHROME A.	6.13e+01
130	5	20.8	460	1	ORC5_DROME	ORIGIN RECOGNITION COM	6.13e+01	203	5	20.8	1165	1	PEX6_PICPA	PEROXISOME BIOSYNTHESI	6.13e+01
131	5	20.8	464	1	BIOA_METJA	ADENOSYLMETHIONINE-8-A	6.13e+01	204	5	20.8	1182	1	METH_MYCLE	5-METHYLTETRAHYDROFOLA	6.13e+01
132	5	20.8	477	1	Y098_MYCGE	HYPOTHETICAL PROTEIN M	6.13e+01	205	5	20.8	1192	1	METH_MYCTU	5-METHYLTETRAHYDROFOLA	6.13e+01
133	5	20.8	480	1	Y066_METJA	HYPOTHETICAL PROTEIN M	6.13e+01	206	5	20.8	1195	1	METH_SYNY3	5-METHYLTETRAHYDROFOLA	6.13e+01
134	5	20.8	480	1	Y066_BOVIN	URIDINE 5'-MONOPHOSPHA	6.13e+01	207	5	20.8	1279	1	APU_THESA	AMYLOPULLULANASE PRECU	6.13e+01
135	5	20.8	480	1	YF05_HUMAN	URIDINE 5'-MONOPHOSPHA	6.13e+01	208	5	20.8	1313	1	VGLM_PTPV	M POLYPROTEIN PRECURSO	6.13e+01
136	5	20.8	487	1	UF05_WANES	FLAVONOL 3-O-GLUCOSYLIT	6.13e+01	209	5	20.8	1316	1	RPOC_MYCTU	DNA-DIRECTED RNA POLYM	6.13e+01
137	5	20.8	490	1	YAM7_SCHPO	HYPOTHETICAL 53.5 KD P	6.13e+01	210	5	20.8	1316	1	RPOC_MYCTU	DNA-DIRECTED RNA POLYM	6.13e+01
138	5	20.8	495	1	VN53_ROTSP	NONSTRUCTURAL RNA-BIND	6.13e+01	211	5	20.8	1398	1	MHP1_YEAST	MAP-HOMOLOGOUS PROTEIN	6.13e+01
139	5	20.8	495	1	VN53_ROTSP	NONSTRUCTURAL RNA-BIND	6.13e+01	212	5	20.8	1473	1	TOP2_ARATH	DNA TOPOISOMERASE II (	6.13e+01
140	5	20.8	508	1	PSBB_SPIOL	PHOTOSYSTEM II P680 CH	6.13e+01	213	5	20.8	1475	1	APU_THETH	AMYLOPULLULANASE PRECU	6.13e+01
141	5	20.8	508	1	PSBB_PINTH	PHOTOSYSTEM II P680 CH	6.13e+01	214	5	20.8	1481	1	APU_THETH	AMYLOPULLULANASE PRECU	6.13e+01
142	5	20.8	508	1	PSBB_SECCB	PHOTOSYSTEM II P680 CH	6.13e+01	215	5	20.8	1505	1	DRQA_DROME	DNA POLYMERASE ALPHA C	6.13e+01
143	5	20.8	508	1	PSBB_ORISA	PHOTOSYSTEM II P680 CH	6.13e+01	216	5	20.8	1589	1	CCZ5_YEAST	CELL DIVISION CONTROL	6.13e+01
144	5	20.8	508	1	PSBB_WHEAT	PHOTOSYSTEM II P680 CH	6.13e+01	217	5	20.8	1603	1	AROL_YEMENI	PENTAFUNCTIONAL AROM P	6.13e+01
145	5	20.8	508	1	PSBB_MAIZE	PHOTOSYSTEM II P680 CH	6.13e+01	218	5	20.8	1615	1	RRPO_TMVR	PUTATIVE RNA-DIRECTED	6.13e+01
146	5	20.8	508	1	PSBB_HORVU	PHOTOSYSTEM II P680 CH	6.13e+01	219	5	20.8	1615	1	RRPO_TMVR	PUTATIVE RNA-DIRECTED	6.13e+01
147	5	20.8	508	1	PSBB_TOBAC	PHOTOSYSTEM II P680 CH	6.13e+01	220	5	20.8	1722	1	YP93_CAEEL	HYPOTHETICAL HELICASE	6.13e+01
148	5	20.8	508	1	PSBB_OENHO	PHOTOSYSTEM II P680 CH	6.13e+01	221	5	20.8	1729	1	RRP5_YEAST	RNA BIOGENESIS PROTEI	6.13e+01
149	5	20.8	509	1	PSBB_POPE	PHOTOSYSTEM II P680 CH	6.13e+01	222	5	20.8	1809	1	TSC2_RAT	TUBERIN (TUBEROUS SCLE	6.13e+01
150	5	20.8	514	1	YUJI_HAEIN	HYPOTHETICAL PROTEIN H	6.13e+01	223	5	20.8	1861	1	APU_THETH	AMYLOPULLULANASE PRECU	6.13e+01
151	5	20.8	545	1	YP17_YEAST	CTP SYNTHASE (EC 6.3.4	6.13e+01	224	5	20.8	1953	1	BN11_YEAST	BN11 PROTEIN (SYNTHETI	6.13e+01
152	5	20.8	551	1	Y098_HAEIN	VACUOLAR PROTEIN SORTI	6.13e+01	225	5	20.8	2167	1	BRM2_YEAST	GTASE ACTIVATING PROT	6.13e+01
153	5	20.8	553	1	P2B1_YEAST	SERINE/THREONINE PROTE	6.13e+01	226	5	20.8	2178	1	CR22_RABIT	BRAIN CALCIUM CHANNEL	6.13e+01
154	5	20.8	555	1	GLPD_BACSU	AEROBIC GLYCEROL-3-PHO	6.13e+01	227	5	20.8	2222	1	CCT1_RAT	NICKEL-SENSITIVE T-TYP	6.13e+01
155	5	20.8	588	1	PP0B_SOLTU	POLYPHENOL OXIDASE B P	6.13e+01	228	5	20.8	2233	1	COAC_YEAST	ACETYL-COA CARBOXYLASE	6.13e+01
156	5	20.8	595	1	VPI_BPCHP	PROTEIN VP1 (ORF1).	6.13e+01	229	5	20.8	2259	1	CB21_RABIT	BRAIN CALCIUM CHANNEL	6.13e+01
157	5	20.8	596	1	SDP_EIMBO	SPOROZOITE DEVELOPMENT	6.13e+01	230	5	20.8	2324	1	COAC_CHICK	ACETYL-COA CARBOXYLASE	6.13e+01
158	5	20.8	604	1	ITAB_PAPCY	PLATELET MEMBRANE GLYC	6.13e+01	231	5	20.8	2345	1	COAC_RAT	ACETYL-COA CARBOXYLASE	6.13e+01
159	5	20.8	605	1	R1B1_RAT	DOLICHYL-DIPHOSPHOOLIG	6.13e+01	232	5	20.8	2346	1	COAC_SHEEP	ACETYL-COA CARBOXYLASE	6.13e+01
160	5	20.8	605	1	AOR_PYRFU	TUNGSTEN-CONTAINING AL	6.13e+01	233	5	20.8	2346	1	COA1_HUMAN	ACETYL-COA CARBOXYLASE	6.13e+01
161	5	20.8	607	1	R1B1_HUMAN	DOLICHYL-DIPHOSPHOOLIG	6.13e+01	234	5	20.8	3084	1	LMAL_MOUSE	LAMININ ALPHA-1 CHAIN	6.13e+01
162	5	20.8	612	1	GIDA_MYCPN	GLUCOSE INHIBITED DIVI	6.13e+01	235	5	20.8	3125	1	POLG_PPVNA	GENOME POLYPROTEIN [CO	6.13e+01
163	5	20.8	614	1	SPAT_BACSU	SUBTILIN TRANSPORT ATP	6.13e+01	236	5	20.8	3140	1	POLG_PPVNA	GENOME POLYPROTEIN [CO	6.13e+01
164	5	20.8	617	1	PY51_PSEAE	PYOCIN S1 (EC 3.1.21.1	6.13e+01	237	5	20.8	3140	1	POLG_PPVSK	GENOME POLYPROTEIN [CO	6.13e+01
165	5	20.8	628	1	SVI_NOSLO	ISOLEUCYL-TRNA SYNTHET	6.13e+01	238	5	20.8	3141	1	POLG_PPVD	GENOME POLYPROTEIN [CO	6.13e+01
166	5	20.8	630	1	Y242_MYCGE	HYPOTHETICAL PROTEIN M	6.13e+01	239	5	20.8	3206	1	POLG_PSBMV	GENOME POLYPROTEIN [CO	6.13e+01
167	5	20.8	632	1	Y242_MYCPN	HYPOTHETICAL PROTEIN M	6.13e+01	240	5	20.8	3759	1	TRX_DROME	TRITHORAX PROTEIN.	6.13e+01
168	5	20.8	641	1	NODO_RHIME	PROBABLE SULFATE ADENY	6.13e+01	241	5	20.8	4092	1	DYHC_YEAST	TRITHORAX HEAVY CHAIN, CY	6.13e+01
169	5	20.8	642	1	SYT_BUCAP	THREONYL-TRNA SYNTHETA	6.13e+01	242	4	16.7	55	1	ATPW_BOVIN	ATP SYNTHASE COUPLING	1.74e+03

243	4	16.7	55	1	ATP8_POLOR	ATP SYNTHASE PROTEIN 8	1.74e+03	316	4	16.7	167	1	OB_BOVIN	LEPTIN PRECURSOR (OBES	1.74e+03
244	4	16.7	57	1	YCUB_CAEEL	HYPOTHETICAL 6.3 KD PR	1.74e+03	317	4	16.7	169	1	YIST_BACSU	HYPOTHETICAL 19.5 KD P	1.74e+03
245	4	16.7	57	1	YCUB_CAEEL	HYPOTHETICAL 6.3 KD PR	1.74e+03	318	4	16.7	169	1	YR7B_ECOTI	HYPOTHETICAL 18.1 KD P	1.74e+03
246	4	16.7	57	1	YOR3_TTV1	HYPOTHETICAL 6.9 KD PR	1.74e+03	319	4	16.7	170	1	Y79B_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
247	4	16.7	57	1	YC5_CAEEL	HYPOTHETICAL 6.3 KD PR	1.74e+03	320	4	16.7	170	1	YCSD_BACSU	HYPOTHETICAL 19.6 KD P	1.74e+03
248	4	16.7	58	1	SAS1_CLOPE	SMALL, ACID-SOLUBLE SP	1.74e+03	321	4	16.7	172	1	Y21B_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
249	4	16.7	62	1	Y12D_BT4	HYPOTHETICAL 7.2 KD PR	1.74e+03	322	4	16.7	173	1	Y15A_ECOTI	INSERTION ELEMENT IS15	1.74e+03
250	4	16.7	68	1	Y975_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	323	4	16.7	178	1	IF3_BUCAP	TRANSLATION INITIATION	1.74e+03
251	4	16.7	71	1	Y47_PORPU	HYPOTHETICAL 8.2 KD PR	1.74e+03	324	4	16.7	179	1	YCF3_ODOSI	HYPOTHETICAL 20.7 KD P	1.74e+03
252	4	16.7	77	1	YOR2_FIVSD	HYPOTHETICAL 9.4 KD PR	1.74e+03	325	4	16.7	180	1	IF3_ECOTI	TRANSLATION INITIATION	1.74e+03
253	4	16.7	78	1	YNP4_CAEEL	HYPOTHETICAL 9.8 KD PR	1.74e+03	326	4	16.7	180	1	IF3_SALTY	TRANSLATION INITIATION	1.74e+03
254	4	16.7	80	1	VSOC_BT4	SMALL OUTER CAPSID PRO	1.74e+03	327	4	16.7	180	1	IF3_KLEPN	TRANSLATION INITIATION	1.74e+03
255	4	16.7	83	1	YFGJ_ECOTI	HYPOTHETICAL 9.2 KD PR	1.74e+03	328	4	16.7	182	1	YLP3_CAEEL	HYPOTHETICAL 20.0 KD P	1.74e+03
256	4	16.7	93	1	YEX6_YEAST	HYPOTHETICAL 10.4 KD P	1.74e+03	329	4	16.7	182	1	YGT2_CHLPS	HYPOTHETICAL 21.0 KD P	1.74e+03
257	4	16.7	93	1	YQ02_MYCTU	HYPOTHETICAL 10.1 KD P	1.74e+03	330	4	16.7	183	1	YH08_YEAST	PUTATIVE MITOCHONDRIAL	1.74e+03
258	4	16.7	96	1	YEA2_SCHPO	VERY HYPOTHETICAL 11.0	1.74e+03	331	4	16.7	183	1	YS4L_CAEEL	PUTATIVE 40S RIBOSOMAL	1.74e+03
259	4	16.7	96	1	Y112_SSV1	HYPOTHETICAL 11.2 KD P	1.74e+03	332	4	16.7	184	1	ZEB2_MAIZE	ZEIN-BETA PRECURSOR (Z	1.74e+03
260	4	16.7	100	1	YJFN_ECOTI	HYPOTHETICAL 11.0 KD P	1.74e+03	333	4	16.7	184	1	YARQ_ACTPL	HYPOTHETICAL 21.6 KD P	1.74e+03
261	4	16.7	102	1	Y10M_MYCPN	HYPOTHETICAL 12.5 KD P	1.74e+03	334	4	16.7	185	1	YGD3_YEAST	HYPOTHETICAL 21.3 KD P	1.74e+03
262	4	16.7	102	1	YQ26_BHP1	HYPOTHETICAL 11.3 KD P	1.74e+03	335	4	16.7	185	1	Y088_MYCTU	HYPOTHETICAL 19.7 KD P	1.74e+03
263	4	16.7	103	1	YDFR_ECOTI	HYPOTHETICAL 12.1 KD P	1.74e+03	336	4	16.7	185	1	YALI_TRYBB	HYPOTHETICAL 22 KD PRO	1.74e+03
264	4	16.7	105	1	Y10B_NPVAC	HYPOTHETICAL 11.8 KD P	1.74e+03	337	4	16.7	186	1	YCEB_ECOTI	20.5 KD PROTEIN IN PYR	1.74e+03
265	4	16.7	106	1	YSMP_IBVU4	PUTATIVE SMALL MEMBRAN	1.74e+03	338	4	16.7	189	1	YQET_BACSU	HYPOTHETICAL 22.2 KD P	1.74e+03
266	4	16.7	108	1	Y108_NPVOP	HYPOTHETICAL 11.7 KD P	1.74e+03	339	4	16.7	191	1	NOT2_YEAST	GENERAL NEGATIVE REGUL	1.74e+03
267	4	16.7	109	1	Y85A_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	340	4	16.7	191	1	Y977_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
268	4	16.7	114	1	YQJ5_CAEEL	HYPOTHETICAL 13.0 KD P	1.74e+03	341	4	16.7	193	1	YK48_YEAST	HYPOTHETICAL 22.1 KD P	1.74e+03
269	4	16.7	114	1	SQDC_DROTO	SUPEROXIDE DISMUTASE I	1.74e+03	342	4	16.7	194	1	Y182_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
270	4	16.7	115	1	YSCN_YEREN	YOP PROTEINS TRANSLOCA	1.74e+03	343	4	16.7	194	1	Y182_ECOTI	HYPOTHETICAL FIMBRIAL -	1.74e+03
271	4	16.7	115	1	YSCM_YERPS	YOP PROTEINS TRANSLOCA	1.74e+03	344	4	16.7	194	1	VP30_ASFB7	PHOSPHOPROTEIN P30.	1.74e+03
272	4	16.7	116	1	YEDM_ECOTI	HYPOTHETICAL 13.4 KD P	1.74e+03	345	4	16.7	194	1	YGF8_ECOTI	HYPOTHETICAL 21.5 KD P	1.74e+03
273	4	16.7	117	1	WNT6_EYATR	WNT-6 PROTEIN (FRAGMEN	1.74e+03	346	4	16.7	195	1	ZN90_HUMAN	ZINC FINGER PROTEIN 90	1.74e+03
274	4	16.7	117	1	WNT6_STRPU	WNT-6 PROTEIN (FRAGMEN	1.74e+03	347	4	16.7	196	1	YK48_YEAST	HYPOTHETICAL 22.1 KD P	1.74e+03
275	4	16.7	119	1	YLT1_ANAVA	HYPOTHETICAL 13.0 KD L	1.74e+03	348	4	16.7	196	1	YKAK_ECOTI	PUTATIVE NAD(P)H OXIDO	1.74e+03
276	4	16.7	120	1	Y104_STRAU	HYPOTHETICAL 14.6 KD P	1.74e+03	349	4	16.7	196	1	VP32_ASFB7	PUTATIVE NAD(P)H OXIDO	1.74e+03
277	4	16.7	121	1	SRB6_YEAST	SUPPRESSOR OF RNA POLY	1.74e+03	350	4	16.7	197	1	YDEQ_BACSU	PUTATIVE NAD(P)H OXIDO	1.74e+03
278	4	16.7	121	1	YPE3_BACSU	HYPOTHETICAL 14.5 KD P	1.74e+03	351	4	16.7	198	1	VS11_ROTBD	MINOR OUTER CAPSID PRO	1.74e+03
279	4	16.7	123	1	VS11_HEVMY	STRUCTURAL PROTEIN 1 (	1.74e+03	352	4	16.7	198	1	RS5_ARCFU	30S RIBOSOMAL PROTEIN	1.74e+03
280	4	16.7	123	1	VS11_HEVBU	STRUCTURAL PROTEIN 1 (	1.74e+03	353	4	16.7	199	1	VTAK_LAMBD	TAIL ASSEMBLY PROTEIN	1.74e+03
281	4	16.7	123	1	VST1_HEVME	STRUCTURAL PROTEIN 1.	1.74e+03	354	4	16.7	199	1	Y599_METJA	HYPOTHETICAL PROTEIN H	1.74e+03
282	4	16.7	123	1	VST1_HEVME	STRUCTURAL PROTEIN 1.	1.74e+03	355	4	16.7	200	1	YE15_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
283	4	16.7	124	1	PA21_BOTJA	PHOSPHOLIPASE A2 (EC 3	1.74e+03	356	4	16.7	203	1	YBM4_YEAST	HYPOTHETICAL 22.6 KD P	1.74e+03
284	4	16.7	129	1	Y42_YEAST	HYPOTHETICAL 14.4 KD P	1.74e+03	357	4	16.7	205	1	YN50_YEAST	HYPOTHETICAL 23.5 KD P	1.74e+03
285	4	16.7	131	1	Y100_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	358	4	16.7	205	1	YVAC_BACSU	HYPOTHETICAL 22.5 KD P	1.74e+03
286	4	16.7	133	1	Y067_CAEEL	HYPOTHETICAL 15.4 KD P	1.74e+03	359	4	16.7	205	1	VNS3_BRD	NONSTRUCTURAL PROTEIN	1.74e+03
287	4	16.7	139	1	YRKE_BACSU	HYPOTHETICAL 16.5 KD P	1.74e+03	360	4	16.7	206	1	RR4_ODOSI	CHLOROPLAST 30S RIBOSO	1.74e+03
288	4	16.7	139	1	Y15K_PMV	HYPOTHETICAL 14.7 KD P	1.74e+03	361	4	16.7	208	1	Y168_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
289	4	16.7	139	1	YPOA_BACSU	HYPOTHETICAL 16.3 KD P	1.74e+03	362	4	16.7	209	1	Y331_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03
290	4	16.7	141	1	RK11_PORPU	CHLOROPLAST 50S RIBOSO	1.74e+03	363	4	16.7	209	1	YCB1_BACSU	HYPOTHETICAL 23.6 KD P	1.74e+03
291	4	16.7	142	1	YB60_YEAST	HYPOTHETICAL 16.3 KD P	1.74e+03	364	4	16.7	210	1	YWAC_BACSU	HYPOTHETICAL 24.6 KD P	1.74e+03
292	4	16.7	143	1	Y4AM_RHISN	HYPOTHETICAL 16.1 KD P	1.74e+03	365	4	16.7	210	1	YT04_MYCTU	HYPOTHETICAL 22.9 KD P	1.74e+03
293	4	16.7	143	1	Y4LB_ENTFA	HYPOTHETICAL 16.4 KD P	1.74e+03	366	4	16.7	210	1	YRL1_METVA	HYPOTHETICAL PROTEIN I	1.74e+03
294	4	16.7	143	1	Y4WC_RHISN	HYPOTHETICAL 16.5 KD P	1.74e+03	367	4	16.7	210	1	YA53_SCHPO	HYPOTHETICAL 24.2 KD P	1.74e+03
295	4	16.7	144	1	DUT_VACCV	DEOXYURIDINE 5'-TRIPHO	1.74e+03	368	4	16.7	212	1	YTFB_ECOTI	HYPOTHETICAL 23.5 KD P	1.74e+03
296	4	16.7	146	1	Y725_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	369	4	16.7	213	1	Y248_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03
297	4	16.7	146	1	YHHA_ECOTI	HYPOTHETICAL 15.6 KD P	1.74e+03	370	4	16.7	213	1	YCFC_ECOTI	HYPOTHETICAL 22.9 KD P	1.74e+03
298	4	16.7	147	1	YSMA_BACSU	HYPOTHETICAL 17.1 KD P	1.74e+03	371	4	16.7	213	1	YKGB_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
299	4	16.7	147	1	YJAB_ECOTI	HYPOTHETICAL 16.4 KD P	1.74e+03	372	4	16.7	216	1	YLAC_SULSO	HYPOTHETICAL 24.4 KD P	1.74e+03
300	4	16.7	150	1	YJWO_YEAST	VERY HYPOTHETICAL 17.0	1.74e+03	373	4	16.7	216	1	YXBE_BACSU	HYPOTHETICAL 23.1 KD P	1.74e+03
301	4	16.7	150	1	YEAL_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	374	4	16.7	217	1	YQBC_BACSU	HYPOTHETICAL 25.3 KD P	1.74e+03
302	4	16.7	151	1	Y105_HUMAN	HYPOTHETICAL PROTEIN K	1.74e+03	375	4	16.7	218	1	YT44_CAEEL	HYPOTHETICAL 25.8 KD P	1.74e+03
303	4	16.7	151	1	Y448_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	376	4	16.7	222	1	YFG0_YEAST	HYPOTHETICAL 25.1 KD P	1.74e+03
304	4	16.7	152	1	VNS2_PVM	NONSTRUCTURAL PROTEIN	1.74e+03	377	4	16.7	222	1	YN74_YEAST	HYPOTHETICAL 25.2 KD P	1.74e+03
305	4	16.7	155	1	YC36_GUITH	HYPOTHETICAL 18.4 KD P	1.74e+03	378	4	16.7	222	1	VP_SVS5	NONSTRUCTURAL PROTEIN	1.74e+03
306	4	16.7	158	1	YACN_BACSU	HYPOTHETICAL 17.1 KD P	1.74e+03	379	4	16.7	222	1	SPSG_BACSU	SPORE COAT POLYSACCHAR	1.74e+03
307	4	16.7	159	1	YGBB_ECOTI	HYPOTHETICAL 16.9 KD P	1.74e+03	380	4	16.7	222	1	YC13_KLEPN	HYPOTHETICAL 24.8 KD P	1.74e+03
308	4	16.7	162	1	YB09_SCHPO	HYPOTHETICAL 18.8 KD P	1.74e+03	381	4	16.7	224	1	YMY5_YEAST	HYPOTHETICAL 24.9 KD P	1.74e+03
309	4	16.7	162	1	Y79A_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	382	4	16.7	226	1	Y24K_HSV4	HYPOTHETICAL 24.6 KD P	1.74e+03
310	4	16.7	164	1	YFDN_ECOTI	HYPOTHETICAL 18.2 KD P	1.74e+03	383	4	16.7	227	1	Y443_METJA	HYPOTHETICAL PROTEIN H	1.74e+03
311	4	16.7	165	1	YC36_PORPU	HYPOTHETICAL 19.2 KD P	1.74e+03	384	4	16.7	227	1	YG24_HAEIN	HYPOTHETICAL PROTEIN M	1.74e+03
312	4	16.7	166	1	YB4B_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	385	4	16.7	228	1	WMS1_CAEEL	HYPOTHETICAL 25.6 KD P	1.74e+03
313	4	16.7	166	1	YB4B_YEAST	HYPOTHETICAL 18.5 KD P	1.74e+03	386	4	16.7	229	1	RR2_ODOSI	CHLOROPLAST 30S RIBOSO	1.74e+03
314	4	16.7	167	1	YHBS_ECOTI	HYPOTHETICAL 18.5 KD P	1.74e+03	387	4	16.7	229	1	RLUB_BACSU	RIBOSOMAL LARGE SUBUNI	1.74e+03
315	4	16.7	167	1	YCX8_ASTLO	HYPOTHETICAL 20.3 KD P	1.74e+03	388	4	16.7	230	1	VP26_NPVOP	P26 PROTEIN.	1.74e+03

389	4	16.7	230	1	YP2A_STAAU	HYPOTHETICAL 26.9 KD P	1.74e+03	462	4	16.7	294	1	YQGI_BACSU	PROBABLE ABC TRANSPORT	1.74e+03
390	4	16.7	231	1	YS29_MYCTU	HYPOTHETICAL 27.1 KD P	1.74e+03	463	4	16.7	295	1	YMR5_YEAST	HYPOTHETICAL 34.5 KD P	1.74e+03
391	4	16.7	231	1	COX2_CAEBL	CYTOCHROME C OXIDASE P	1.74e+03	464	4	16.7	295	1	YFLF_CAEBL	HYPOTHETICAL PROTEIN I	1.74e+03
392	4	16.7	231	1	Y18T_MYCTU	HYPOTHETICAL 24.0 KD P	1.74e+03	465	4	16.7	296	1	Y84T_CAEBL	HYPOTHETICAL 33.5 KD P	1.74e+03
393	4	16.7	234	1	YNK0_YEAST	HYPOTHETICAL 27.0 KD P	1.74e+03	466	4	16.7	297	1	YNFL_ECOLI	HYPOTHETICAL TRANSCRIP	1.74e+03
394	4	16.7	234	1	YGS5_ECOLI	HYPOTHETICAL 25.8 KD P	1.74e+03	467	4	16.7	298	1	YNT6_YEAST	HYPOTHETICAL 34.3 KD P	1.74e+03
395	4	16.7	235	1	YD3Z_ECOLI	HYPOTHETICAL 26.2 KD P	1.74e+03	468	4	16.7	298	1	Y1HV_ECOLI	HYPOTHETICAL SUGAR KIN	1.74e+03
396	4	16.7	235	1	RR2_GUTH	CHLOROPLAST 30S RIBOSO	1.74e+03	469	4	16.7	299	1	SP11_CAEBL	SPERMATOCYTE PROTEIN S	1.74e+03
397	4	16.7	236	1	YKA9_YEAST	HYPOTHETICAL 27.1 KD P	1.74e+03	470	4	16.7	301	1	YF04_YEAST	HYPOTHETICAL 33.5 KD P	1.74e+03
398	4	16.7	236	1	Y110_MYCGE	HYPOTHETICAL PROTEIN M	1.74e+03	471	4	16.7	301	1	YR39_MYCTU	HYPOTHETICAL 32.9 KD P	1.74e+03
399	4	16.7	237	1	YQ04_CAEBL	HYPOTHETICAL 26.2 KD P	1.74e+03	472	4	16.7	302	1	RS3_HALHA	30S RIBOSOMAL PROTEIN	1.74e+03
400	4	16.7	238	1	Y1GB_ECOLI	HYPOTHETICAL 27.1 KD P	1.74e+03	473	4	16.7	302	1	COX2_RHOSH	CYTOCHROME C OXIDASE P	1.74e+03
401	4	16.7	239	1	RY5_ODOSI	HYPOTHETICAL 26.1 KD P	1.74e+03	474	4	16.7	303	1	SC14_YEAST	SEC14 CYTOSOLIC FACTOR	1.74e+03
402	4	16.7	239	1	Y027_MYCPN	CHLOROPLAST 50S RIBOSO	1.74e+03	475	4	16.7	304	1	RS3_HALMA	30S RIBOSOMAL PROTEIN	1.74e+03
403	4	16.7	240	1	VT4_CAPVI	HYPOTHETICAL PROTEIN M	1.74e+03	476	4	16.7	305	1	YQBD_BACSU	HYPOTHETICAL 34.6 KD P	1.74e+03
404	4	16.7	240	1	YTA_CAPVK	T4 PROTEIN	1.74e+03	477	4	16.7	305	1	YQBB_BACSU	HYPOTHETICAL 34.9 KD P	1.74e+03
405	4	16.7	241	1	YJHA_ECOLI	HYPOTHETICAL 28.3 KD P	1.74e+03	478	4	16.7	305	1	YK61_CAEBL	PUTATIVE CUTICLE COLLA	1.74e+03
406	4	16.7	241	1	RR2_PORPU	CHLOROPLAST 30S RIBOSO	1.74e+03	479	4	16.7	306	1	BLAC_STRFR	BETA-LACTAMASE PRECURS	1.74e+03
407	4	16.7	243	1	Y124_NPVOP	HYPOTHETICAL 26.5 KD P	1.74e+03	480	4	16.7	306	1	MMGL_RAT	MACROPHAGE ASIALOGLYCO	1.74e+03
408	4	16.7	243	1	Y027_CAEBL	HYPOTHETICAL 26.5 KD P	1.74e+03	481	4	16.7	307	1	YD44_SCHPO	PUTATIVE SERINE/THREON	1.74e+03
409	4	16.7	243	1	RS3_HUMAN	40S RIBOSOMAL PROTEIN	1.74e+03	482	4	16.7	307	1	Y608_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
410	4	16.7	244	1	YU58_CAEBL	HYPOTHETICAL 27.2 KD P	1.74e+03	483	4	16.7	310	1	YGLA_SYN2	HYPOTHETICAL 34.1 KD P	1.74e+03
411	4	16.7	244	1	YF1H_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	484	4	16.7	310	1	YV23_MYCTU	HYPOTHETICAL 33.1 KD P	1.74e+03
412	4	16.7	245	1	YBV7_YEAST	HYPOTHETICAL 28.1 KD P	1.74e+03	485	4	16.7	313	1	YFM4_MYCOL	HYPOTHETICAL PROTEIN I	1.74e+03
413	4	16.7	245	1	Y01A_MYCTU	HYPOTHETICAL 26.9 KD P	1.74e+03	486	4	16.7	315	1	YOH1_ECOLI	HYPOTHETICAL 35.2 KD P	1.74e+03
414	4	16.7	245	1	Y1JG_ECOLI	HYPOTHETICAL 26.3 KD P	1.74e+03	487	4	16.7	315	1	ASPG_ARATH	L-ASPARAGINASE (EC 3.5	1.74e+03
415	4	16.7	246	1	RS3_DROME	40S RIBOSOMAL PROTEIN	1.74e+03	488	4	16.7	316	1	Y184_MYCGE	HYPOTHETICAL PROTEIN M	1.74e+03
416	4	16.7	246	1	YD53_SCHPO	40S RIBOSOMAL PROTEIN	1.74e+03	489	4	16.7	317	1	Y184_MYCGE	HYPOTHETICAL ADENINE-S	1.74e+03
417	4	16.7	247	1	FGFE_HUMAN	FIBROBLAST GROWTH FACT	1.74e+03	490	4	16.7	317	1	Y600_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
418	4	16.7	247	1	ZO20_XENLA	OOCYTE ZINC FINGER PRO	1.74e+03	491	4	16.7	318	1	Y940_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
419	4	16.7	247	1	YGF5_ECOLI	HYPOTHETICAL OXIDOREDU	1.74e+03	492	4	16.7	319	1	YKM1_CAEBL	HYPOTHETICAL 35.5 KD P	1.74e+03
420	4	16.7	254	1	Y194_SYN3	HYPOTHETICAL 27.8 KD P	1.74e+03	493	4	16.7	319	1	Y018_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
421	4	16.7	255	1	YN04_PARDE	HYPOTHETICAL 26.9 KD P	1.74e+03	494	4	16.7	320	1	YB1B_ECOLI	HYPOTHETICAL 35.0 KD P	1.74e+03
422	4	16.7	257	1	Y08C_MYCTU	HYPOTHETICAL 26.4 KD P	1.74e+03	495	4	16.7	320	1	Y381_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
423	4	16.7	257	1	PSTB_ECOLI	PHOSPHATE TRANSPORT AT	1.74e+03	496	4	16.7	322	1	YJ96_CAEBL	HYPOTHETICAL 36.9 KD P	1.74e+03
424	4	16.7	258	1	YGBM_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	497	4	16.7	322	1	YFX4_CAEBL	HYPOTHETICAL 36.5 KD P	1.74e+03
425	4	16.7	259	1	Y128_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	498	4	16.7	322	1	SC13_HUMAN	SEC13-RELATED PROTEIN	1.74e+03
426	4	16.7	261	1	Y152_MYCTU	HYPOTHETICAL 26.9 KD P	1.74e+03	499	4	16.7	324	1	Y700_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
427	4	16.7	261	1	YV23_MYCLE	HYPOTHETICAL 27.9 KD P	1.74e+03	500	4	16.7	324	1	CATV_NPVCF	VIRAL CATHEPSIN (EC 3.	1.74e+03
428	4	16.7	262	1	YGS6_YEAST	HYPOTHETICAL 29.4 KD P	1.74e+03	501	4	16.7	324	1	YAGC_SCHPO	HYPOTHETICAL 35.8 KD	1.74e+03
429	4	16.7	262	1	Y4UP_RHISN	HYPOTHETICAL 29.5 KD P	1.74e+03	502	4	16.7	325	1	V72_SFVKA	TUMOR NECROSIS FACTOR	1.74e+03
430	4	16.7	263	1	SSA1_MACFA	3-OXO-5-ALPHA-STEROID	1.74e+03	503	4	16.7	325	1	YXDK_BACSU	HYPOTHETICAL SENSOR-LI	1.74e+03
431	4	16.7	266	1	YDEB_ECOLI	HYPOTHETICAL 28.7 KD P	1.74e+03	504	4	16.7	325	1	Y05F_BPT4	HYPOTHETICAL 36.7 KD P	1.74e+03
432	4	16.7	267	1	RS2_DROME	40S RIBOSOMAL PROTEIN	1.74e+03	505	4	16.7	325	1	Y0A8_MYCTU	HYPOTHETICAL 33.1 KD P	1.74e+03
433	4	16.7	269	1	Y374_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	506	4	16.7	326	1	YDJG_ECOLI	HYPOTHETICAL OXIDOREDU	1.74e+03
434	4	16.7	269	1	YHT4_YEAST	HYPOTHETICAL 30.6 KD P	1.74e+03	507	4	16.7	326	1	YKX2_CAEBL	HYPOTHETICAL 34.6 KD P	1.74e+03
435	4	16.7	272	1	YCR3_ERWHE	HYPOTHETICAL 29.9 KD P	1.74e+03	508	4	16.7	326	1	YV40_EBOV	MATRIX PROTEIN VP40.	1.74e+03
436	4	16.7	273	1	YGF0_YEAST	HYPOTHETICAL 30.8 KD P	1.74e+03	509	4	16.7	327	1	YK79_SYN3	HYPOTHETICAL 36.0 KD P	1.74e+03
437	4	16.7	274	1	RS3_MYCBO	HYPOTHETICAL 31.6 KD P	1.74e+03	510	4	16.7	327	1	YD96_SCHPO	HYPOTHETICAL 38.1 KD P	1.74e+03
438	4	16.7	274	1	Y939_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	511	4	16.7	328	1	Y052_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
439	4	16.7	276	1	Y842_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	512	4	16.7	329	1	PSTB_MYCPN	PHOSPHATE TRANSPORT AT	1.74e+03
440	4	16.7	276	1	Y578_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	513	4	16.7	329	1	YK88_YEAST	HYPOTHETICAL 37.1 KD P	1.74e+03
441	4	16.7	276	1	Y110_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	514	4	16.7	330	1	YXA0_BACSU	HYPOTHETICAL 37.5 KD P	1.74e+03
442	4	16.7	278	1	YKDY_BACSU	HYPOTHETICAL PROTEIN M	1.74e+03	515	4	16.7	330	1	PER2_ARAHY	CATIONIC PEROXIDASE 2	1.74e+03
443	4	16.7	279	1	YX46_CAEBL	PHAGE-LIKE ELEMENT PBS	1.74e+03	516	4	16.7	330	1	Y677_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
444	4	16.7	279	1	VRXA_LAMBD	HYPOTHETICAL 32.5 KD P	1.74e+03	517	4	16.7	333	1	Y808_METJA	PROBABLE ABC TRANSPORT	1.74e+03
445	4	16.7	279	1	FRHB_WETTH	COENZYME F420 HYDROGEN	1.74e+03	518	4	16.7	333	1	Y4MJ_RHISN	PROBABLE ABC TRANSPORT	1.74e+03
446	4	16.7	280	1	FRHB_WETTH	COENZYME F420 HYDROGEN	1.74e+03	519	4	16.7	334	1	VPRT_HUMAN	RETROVIRUS-RELATED PRO	1.74e+03
447	4	16.7	283	1	Y43C_MYCPN	HYPOTHETICAL 31.2 KD P	1.74e+03	520	4	16.7	334	1	YXJG_BACSU	HYPOTHETICAL 38.0 KD P	1.74e+03
448	4	16.7	283	1	YK26_YEAST	HYPOTHETICAL 31.2 KD P	1.74e+03	521	4	16.7	337	1	YU17_YEAST	HYPOTHETICAL 37.9 KD P	1.74e+03
449	4	16.7	285	1	Y19J_MYCTU	HYPOTHETICAL 29.7 KD P	1.74e+03	522	4	16.7	338	1	YEY6_YEAST	HYPOTHETICAL 38.2 KD P	1.74e+03
450	4	16.7	285	1	Y219_MYCTU	HYPOTHETICAL 31.2 KD P	1.74e+03	523	4	16.7	339	1	YJ99_YEAST	HYPOTHETICAL 39.0 KD P	1.74e+03
451	4	16.7	285	1	RS2_ARATH	40S RIBOSOMAL PROTEIN	1.74e+03	524	4	16.7	339	1	YF73_SCHPO	PUTATIVE 38.2 KD PHOS	1.74e+03
452	4	16.7	285	1	YHMF_METFE	HYPOTHETICAL 32.2 KD P	1.74e+03	525	4	16.7	340	1	YKXH_BACSU	HYPOTHETICAL 38.3 KD P	1.74e+03
453	4	16.7	288	1	YLN9_CAEBL	HYPOTHETICAL 32.9 KD P	1.74e+03	526	4	16.7	340	1	Y05A_BPT4	HYPOTHETICAL 39.7 KD P	1.74e+03
454	4	16.7	288	1	YQD5_CAEBL	HYPOTHETICAL 33.2 KD P	1.74e+03	527	4	16.7	342	1	YK22_YEAST	HYPOTHETICAL 38.6 KD P	1.74e+03
455	4	16.7	288	1	YHCS_HAEIN	HYPOTHETICAL TRANSCRIP	1.74e+03	528	4	16.7	344	1	YK58_ARCFU	HYPOTHETICAL PROTEIN A	1.74e+03
456	4	16.7	289	1	YGV2_YEAST	HYPOTHETICAL 33.6 KD P	1.74e+03	529	4	16.7	346	1	YA0D_SCHPO	HYPOTHETICAL 39.1 KD P	1.74e+03
457	4	16.7	290	1	Y43A_MYCPN	HYPOTHETICAL LIPOPROTE	1.74e+03	530	4	16.7	346	1	LUXA_PHOPO	ALKANAL MONOOXYGENASE	1.74e+03
458	4	16.7	290	1	YMA3_YEAST	HYPOTHETICAL 34.0 KD P	1.74e+03	531	4	16.7	348	1	FML1_MACMU	FMLP-RELATED RECEPTOR	1.74e+03
459	4	16.7	290	1	BUDR_KLETE	BUD OPERON TRANSCRIPTI	1.74e+03	532	4	16.7	350	1	VCEA_ECOLI	HYPOTHETICAL 40.0 KD P	1.74e+03
460	4	16.7	293	1	YQUS_BACSU	HYPOTHETICAL 33.6 KD P	1.74e+03	533	4	16.7	350	1	XYLA_PSEPU	XYLENE MONOOXYGENASE	1.74e+03
461	4	16.7	293	1	RS2_HUMAN	40S RIBOSOMAL PROTEIN	1.74e+03	534	4	16.7	350	1	RUVB_TREPA	HOLLIDAY JUNCTION DNA	1.74e+03

535	4	16.7	352	1	SMP_SERVA	EXTRACELLULAR MINOR ME	1.74e+03	608	4	16.7	413	1	YAA7_SCHPO	HYPOTHETICAL 47.3 KD P	1.74e+03
536	4	16.7	353	1	YQGU_BACSU	HYPOTHETICAL 40.3 KD L	1.74e+03	609	4	16.7	413	1	YP71_MYCMY	HYPOTHETICAL 47.4 KD P	1.74e+03
537	4	16.7	354	1	YHL4_YEAST	HYPOTHETICAL 41.1 KD P	1.74e+03	610	4	16.7	413	1	YIHS_ECOLI	HYPOTHETICAL 47.4 KD P	1.74e+03
538	4	16.7	355	1	YHL4_SYNY3	HYPOTHETICAL 39.3 KD P	1.74e+03	611	4	16.7	414	1	Y694_METJA	HYPOTHETICAL 47.8 KD P	1.74e+03
539	4	16.7	356	1	YD6B_SCHPO	HYPOTHETICAL 42.2 KD P	1.74e+03	612	4	16.7	414	1	Y444_GUITH	HYPOTHETICAL 47.8 KD P	1.74e+03
540	4	16.7	356	1	VU8_HSV6U	U8 PROTEIN.	1.74e+03	613	4	16.7	414	1	YS25_CAEEL	HYPOTHETICAL 45.3 KD P	1.74e+03
541	4	16.7	356	1	YJN1_YEAST	HYPOTHETICAL 41.5 KD P	1.74e+03	614	4	16.7	418	1	YB01_FOWPM	HYPOTHETICAL 45.3 KD P	1.74e+03
542	4	16.7	358	1	YDGO_HAEIN	HYPOTHETICAL 43.0 KD P	1.74e+03	615	4	16.7	419	1	RNBP_RAT	RENIN-BINDING PROTEIN	1.74e+03
543	4	16.7	358	1	VP43_NPVAC	43 KD PROTEIN.	1.74e+03	616	4	16.7	419	1	P47B_CANBO	PEROXISOMAL MEMBRANE P	1.74e+03
544	4	16.7	363	1	YF18_YEAST	HYPOTHETICAL 41.0 KD P	1.74e+03	617	4	16.7	419	1	YMD2_SCHPO	HYPOTHETICAL 47.3 KD P	1.74e+03
545	4	16.7	363	1	YF18_YEAST	HYPOTHETICAL 41.0 KD P	1.74e+03	618	4	16.7	419	1	YDEA_SCHPO	HYPOTHETICAL 48.1 KD P	1.74e+03
546	4	16.7	363	1	PGLR_ECOLI	ABRB PROTEIN.	1.74e+03	619	4	16.7	423	1	YJ54_YEAST	HYPOTHETICAL 49.5 KD P	1.74e+03
547	4	16.7	363	1	PGLR_APPPA	POLYGLUTAMINASE PREC	1.74e+03	620	4	16.7	423	1	BRAC_BRARE	BRACHYURY PROTEIN HOMO	1.74e+03
548	4	16.7	363	1	WNT5_BARE	WNT-5 PROTEIN PRECURSO	1.74e+03	621	4	16.7	424	1	Y146_MYCPN	HYPOTHETICAL 49.5 KD P	1.74e+03
549	4	16.7	366	1	Y433_ARCFU	HYPOTHETICAL 41.0 KD P	1.74e+03	622	4	16.7	424	1	ZFPI_MOUSE	ZINC FINGER PROTEIN 2F	1.74e+03
550	4	16.7	366	1	YQEH_BACSU	HYPOTHETICAL 41.0 KD P	1.74e+03	623	4	16.7	425	1	YHAO_ECOLI	HYPOTHETICAL 46.6 KD P	1.74e+03
551	4	16.7	366	1	Y976_METJA	HYPOTHETICAL 40.7 KD P	1.74e+03	624	4	16.7	425	1	YGGP_ECOLI	HYPOTHETICAL 45.8 KD P	1.74e+03
552	4	16.7	366	1	YHP7_YEAST	HYPOTHETICAL 40.7 KD P	1.74e+03	625	4	16.7	426	1	YAJ6_SCHPO	HYPOTHETICAL 49.3 KD P	1.74e+03
553	4	16.7	367	1	Y024_MYCGE	PROBABLE GTP-BINDING P	1.74e+03	626	4	16.7	426	1	VP8_RGDV	OUTER CAPSID PROTEIN P	1.74e+03
554	4	16.7	368	1	YM39_YEAST	PUTATIVE MITOCHONDRIAL	1.74e+03	627	4	16.7	427	1	Y4FK_RHISN	PUTATIVE TRANSCRIPTION	1.74e+03
555	4	16.7	368	1	YIU5_YEAST	HYPOTHETICAL 42.8 KD P	1.74e+03	628	4	16.7	427	1	NGFR_HUMAN	LOW-AFFINITY NERVE GRO	1.74e+03
556	4	16.7	369	1	YI05_YEAST	HYPOTHETICAL 41.0 KD P	1.74e+03	629	4	16.7	428	1	YQPI_CAEEL	HYPOTHETICAL 47.6 KD P	1.74e+03
557	4	16.7	369	1	Y264_SYNY3	HOLLIDAY JUNCTION DNA	1.74e+03	630	4	16.7	428	1	YMI4_CAEEL	HYPOTHETICAL 48.3 KD P	1.74e+03
558	4	16.7	370	1	YGP9_YEAST	HYPOTHETICAL 41.6 KD P	1.74e+03	631	4	16.7	431	1	YQAT_BACSU	HYPOTHETICAL 50.9 KD P	1.74e+03
559	4	16.7	370	1	VP40_BDV	40 KD PROTEIN.	1.74e+03	632	4	16.7	431	1	YMA9_YEAST	HYPOTHETICAL 50.3 KD P	1.74e+03
560	4	16.7	371	1	WNT1_XENLA	XWNT-1 PROTEIN PRECURS	1.74e+03	633	4	16.7	432	1	Y4IM_RHISN	HYPOTHETICAL 47.1 KD P	1.74e+03
561	4	16.7	373	1	YCO8_KLEPN	HYPOTHETICAL 42.6 KD P	1.74e+03	634	4	16.7	433	1	XTMB_BACSU	PBSX PHASE TERMINASE L	1.74e+03
562	4	16.7	373	1	YV32_MYCTU	HYPOTHETICAL 39.5 KD P	1.74e+03	635	4	16.7	433	1	RPD3_YEAST	HISTONE DEACETYLASE RP	1.74e+03
563	4	16.7	374	1	RT04_ACACA	MITOCHONDRIAL 41.0 KD P	1.74e+03	636	4	16.7	434	1	Y181_MYCPN	HYPOTHETICAL 47.6 KD P	1.74e+03
564	4	16.7	375	1	YK05_CAEEL	HYPOTHETICAL 41.5 KD P	1.74e+03	637	4	16.7	434	1	YSX4_CAEEL	HYPOTHETICAL 50.4 KD P	1.74e+03
565	4	16.7	375	1	NUEM_NEUCR	NADH-UBIQUINONE OXIDOR	1.74e+03	638	4	16.7	434	1	YUGS_BACSU	HYPOTHETICAL 49.5 KD P	1.74e+03
566	4	16.7	376	1	NIR_ALCPA	COPPER-CONTAINING NITR	1.74e+03	639	4	16.7	434	1	YLP4_CAEEL	HYPOTHETICAL 49.1 KD P	1.74e+03
567	4	16.7	376	1	YHL2_YEAST	HYPOTHETICAL 42.5 KD P	1.74e+03	640	4	16.7	435	1	Y413_ARATH	HYPOTHETICAL 48.8 KD P	1.74e+03
568	4	16.7	378	1	YMP6_CAEEL	HYPOTHETICAL 43.0 KD P	1.74e+03	641	4	16.7	439	1	Z028_XENLA	OOCYTE ZINC FINGER PRO	1.74e+03
569	4	16.7	378	1	YIBH_ECOLI	HYPOTHETICAL 42.0 KD P	1.74e+03	642	4	16.7	440	1	YB54_HAEIN	HYPOTHETICAL SYMPORTER	1.74e+03
570	4	16.7	380	1	YPG2_ZYMMO	HYPOTHETICAL 43.3 KD P	1.74e+03	643	4	16.7	441	1	YD40_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
571	4	16.7	380	1	YKAD_ECOLI	HYPOTHETICAL 41.4 KD P	1.74e+03	644	4	16.7	441	1	YLIQ_ECOLI	HYPOTHETICAL 49.6 KD P	1.74e+03
572	4	16.7	382	1	YXIM_BACSU	HYPOTHETICAL 41.8 KD P	1.74e+03	645	4	16.7	441	1	YK40_PIRHO	HYPOTHETICAL ABC TRANS	1.74e+03
573	4	16.7	383	1	NTRB_RHLIP	NITROGEN REGULATION PR	1.74e+03	646	4	16.7	442	1	RRG2_XENLA	RETINOIC ACID RECEPTOR	1.74e+03
574	4	16.7	383	1	XYLR_STAXY	XYLOSE REPRESSOR.	1.74e+03	647	4	16.7	444	1	YHPD_BACSU	RETINOIC ACID RECEPTOR	1.74e+03
575	4	16.7	383	1	Y776_METJA	HYPOTHETICAL 43.8 KD P	1.74e+03	648	4	16.7	445	1	YHPD_BACSU	HYPOTHETICAL 49.9 KD P	1.74e+03
576	4	16.7	384	1	YAGA_ECOLI	HYPOTHETICAL 43.8 KD P	1.74e+03	649	4	16.7	446	1	Y612_METJA	HYPOTHETICAL 47.8 KD P	1.74e+03
577	4	16.7	384	1	YFGB_ECOLI	HYPOTHETICAL 43.1 KD P	1.74e+03	650	4	16.7	446	1	Y164_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
578	4	16.7	384	1	YXT2_YEAST	HYPOTHETICAL COB INTRO	1.74e+03	651	4	16.7	447	1	RRG2_MOUSE	RETINOIC ACID RECEPTOR	1.74e+03
579	4	16.7	385	1	PELE_ERMCH	PECTATE LYASE E PRECUR	1.74e+03	652	4	16.7	448	1	RRB2_HUMAN	RETINOIC ACID RECEPTOR	1.74e+03
580	4	16.7	385	1	YD16_SCHPO	HYPOTHETICAL 41.0 KD P	1.74e+03	653	4	16.7	449	1	APG_BRANA	ANTER-SPECIFIC PROLINE	1.74e+03
581	4	16.7	386	1	YAAH_BACSU	HYPOTHETICAL 43.8 KD P	1.74e+03	654	4	16.7	451	1	NTPI_ENTHR	V-TYPE SODIUM ATP SYNT	1.74e+03
582	4	16.7	387	1	VP48_NPVAC	P48 PROTEIN (P45).	1.74e+03	655	4	16.7	451	1	Y4V1_RHISN	PROBABLE ATP SYNTHASE	1.74e+03
583	4	16.7	388	1	YGJO_ECOLI	HYPOTHETICAL 43.4 KD P	1.74e+03	656	4	16.7	451	1	Y041_METJA	HYPOTHETICAL PROTEIN M	1.74e+03
584	4	16.7	389	1	Y293_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	657	4	16.7	455	1	YMFN_ECOLI	HYPOTHETICAL 50.9 KD P	1.74e+03
585	4	16.7	393	1	YHFC_ECOLI	HYPOTHETICAL 43.2 KD P	1.74e+03	658	4	16.7	455	1	YS11_REOVD	SIGMA 1 PROTEIN PRECUR	1.74e+03
586	4	16.7	395	1	Y443_MYCGE	HYPOTHETICAL PROTEIN M	1.74e+03	659	4	16.7	456	1	YC13_ASTLO	HYPOTHETICAL 54.4 KD P	1.74e+03
587	4	16.7	395	1	NEUA_STRAG	ACYLNEURAMINATE CYTIDY	1.74e+03	660	4	16.7	457	1	NTRC_RHOCA	NITROGEN REGULATION PR	1.74e+03
588	4	16.7	397	1	Y048_MYCTU	HYPOTHETICAL 42.4 KD P	1.74e+03	661	4	16.7	458	1	YKH3_CAEEL	HYPOTHETICAL ACETYLCHO	1.74e+03
589	4	16.7	398	1	Y665_METJA	PUTATIVE MOLYBDOPTEIN	1.74e+03	662	4	16.7	459	1	XDJ1_YEAST	XDJ1 PROTEIN.	1.74e+03
590	4	16.7	399	1	P2X1_RAT	P2X PURINOCCEPTOR 1 (AT	1.74e+03	663	4	16.7	459	1	FLHF_HELPY	FLAGELLAR BIOSYNTHESIS	1.74e+03
591	4	16.7	399	1	YCFU_ECOLI	HYPOTHETICAL 43.3 KD P	1.74e+03	664	4	16.7	462	1	YS11_REOJV	SIGMA 1 PROTEIN PRECUR	1.74e+03
592	4	16.7	401	1	YHIN_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03	665	4	16.7	462	1	RRAL_HUMAN	RETINOIC ACID RECEPTOR	1.74e+03
593	4	16.7	401	1	Y443_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	666	4	16.7	462	1	Y154_SYNY3	HYPOTHETICAL 51.0 KD P	1.74e+03
594	4	16.7	402	1	YIN2_STRAM	HYPOTHETICAL 44.6 KD P	1.74e+03	667	4	16.7	463	1	YC58_YEAST	HYPOTHETICAL SP52-LIKE	1.74e+03
595	4	16.7	402	1	YJT3_RHISN	HYPOTHETICAL HYDROLASE	1.74e+03	668	4	16.7	464	1	WCAM_ECOLI	COLANIC ACID BIOSYNTH	1.74e+03
596	4	16.7	402	1	YJTT_YEAST	HYPOTHETICAL 45.9 KD P	1.74e+03	669	4	16.7	465	1	YHJA_ECOLI	PROBABLE CYTOCHROME C	1.74e+03
597	4	16.7	402	1	YXAH_BACSU	HYPOTHETICAL 46.2 KD P	1.74e+03	670	4	16.7	465	1	YCJX_ECOLI	HYPOTHETICAL 52.6 KD P	1.74e+03
598	4	16.7	402	1	YQN3_CAEEL	PROBABLE PSEUDOURIDYLA	1.74e+03	671	4	16.7	466	1	Y065_MYCGE	HYPOTHETICAL ABC TRANS	1.74e+03
599	4	16.7	404	1	YJHT_ECOLI	HYPOTHETICAL 43.7 KD P	1.74e+03	672	4	16.7	469	1	YBS8_YEAST	HYPOTHETICAL 48.3 KD P	1.74e+03
600	4	16.7	406	1	YCX9_EUGGR	HYPOTHETICAL 45.2 KD P	1.74e+03	673	4	16.7	470	1	BFRL_YEAST	NUCLEAR SEGREGATION PR	1.74e+03
601	4	16.7	406	1	YMX2_CAEEL	HYPOTHETICAL 47.7 KD P	1.74e+03	674	4	16.7	470	1	YCJX_HAEIN	HYPOTHETICAL PROTEIN H	1.74e+03
602	4	16.7	407	1	YXN1_METTH	PUTATIVE AMMONIUM TRAN	1.74e+03	675	4	16.7	470	1	YICO_ECOLI	HYPOTHETICAL 49.9 KD P	1.74e+03
603	4	16.7	407	1	YXN1_BACST	ENDO-1,4-BETA-XYLANASE	1.74e+03	676	4	16.7	472	1	NU49_YEAST	NUCLEOPORIN NUP49/NSP4	1.74e+03
604	4	16.7	408	1	AMPT_THETH	AMINOPEPTIDASE T (EC 3	1.74e+03	677	4	16.7	474	1	YSM5_TRYBB	VARIANT SURFACE GLYCOP	1.74e+03
605	4	16.7	408	1	Y889_SYNY3	HYPOTHETICAL 45.0 KD P	1.74e+03	678	4	16.7	474	1	YCSC_BACSU	PUTATIVE L-LACTATE PER	1.74e+03
606	4	16.7	409	1	YV12_MYCLE	HYPOTHETICAL 43.5 KD P	1.74e+03	679	4	16.7	475	1	YM61_YEAST	HYPOTHETICAL 55.3 KD P	1.74e+03
607	4	16.7	412	1	ASSY_BOVIN	ARGININOSUCCINATE SYNT	1.74e+03	680	4	16.7	475	1	RP54_THIFE	RNA POLYMERASE SIGMA-5	1.74e+03

681	4	16.7	479	1	XYLP_LACPE	PUTATIVE XYLOSE-PROTON	1.74e+03	754	4	16.7	569	1	YAR3_SCHPO	HYPOTHETICAL 61.1 KD P	1.74e+03
682	4	16.7	481	1	YABC_SCHPO	HYPOTHETICAL 54.4 KD P	755	4	16.7	572	1	DPY2_MOUSE	DIHYDROPYRIMIDINASE RE	1.74e+03	
683	4	16.7	481	1	YAB8_SCHPO	PUTATIVE ATP-DEPENDENT	756	4	16.7	574	1	IPM4_SHIFL	65.4 KD ANTIGEN	1.74e+03	
684	4	16.7	481	1	YRDD_CABEL	HYPOTHETICAL 53.4 KD P	757	4	16.7	574	1	XJH1_YEAST	HYPOTHETICAL 65.6 KD P	1.74e+03	
685	4	16.7	483	1	NIFE_METMP	NITROGENASE IRON-MOLYB	758	4	16.7	575	1	SMF1_YEAST	TRANSPORTER PROTEIN SM	1.74e+03	
686	4	16.7	484	1	YOD0_YEAST	HYPOTHETICAL 51.9 KD P	759	4	16.7	575	1	RM62_DROME	PUTATIVE ATP-DEPENDENT	1.74e+03	
687	4	16.7	486	1	SAHH_TRIVA	ADENOSYLHOMOCYSTEINASE	760	4	16.7	576	1	XYL2_YEAST	HYPOTHETICAL 64.0 KD P	1.74e+03	
688	4	16.7	486	1	YAM8_SCHPO	HYPOTHETICAL 53.9 KD P	761	4	16.7	577	1	SYD_HELPY	ASPARTYL-TRNA SYNTHETA	1.74e+03	
689	4	16.7	487	1	YC24_PORPU	HYPOTHETICAL 54.4 KD P	762	4	16.7	578	1	YBHE_ECOLI	HYPOTHETICAL ABC TRANS	1.74e+03	
690	4	16.7	489	1	YK27_CABEL	HYPOTHETICAL 54.9 KD P	763	4	16.7	578	1	YHA2_EIKCO	HYPOTHETICAL 66.3 KD P	1.74e+03	
691	4	16.7	489	1	YN21_CABEL	PUTATIVE ATP-DEPENDENT	764	4	16.7	578	1	YDEM_CABEL	HYPOTHETICAL 65.5 KD T	1.74e+03	
692	4	16.7	490	1	YIHF_ECOLI	HYPOTHETICAL 54.2 KD P	765	4	16.7	580	1	ATPUL_YEAST	ATP13 PROTEIN PRECURSOR	1.74e+03	
693	4	16.7	492	1	YSM1_TRYBB	VARIANT SURFACE GLYCOP	766	4	16.7	582	1	Y089_MYCTU	HYPOTHETICAL ABC TRANS	1.74e+03	
694	4	16.7	495	1	VPE_SOYBN	VACUOLAR PROCESSING EN	767	4	16.7	586	1	RT67_ECOLI	RNA-DIRECTED DNA POLYM	1.74e+03	
695	4	16.7	497	1	YP25_CABEL	HYPOTHETICAL 56.2 KD P	768	4	16.7	587	1	YNV8_YEAST	HYPOTHETICAL 66.5 KD P	1.74e+03	
696	4	16.7	498	1	YIEN_ECOLI	HYPOTHETICAL 56.4 KD P	769	4	16.7	589	1	STI1_YEAST	HEAT SHOCK PROTEIN STI	1.74e+03	
697	4	16.7	498	1	YCJX_ECOLI	PUTATIVE GLUTAMINE SYN	770	4	16.7	590	1	YKTA_CABEL	HYPOTHETICAL 65.5 KD P	1.74e+03	
698	4	16.7	498	1	YJUI_YEAST	HYPOTHETICAL 56.4 KD P	771	4	16.7	590	1	VPP_BPP2	TERMINASE, ATPASE SUBU	1.74e+03	
699	4	16.7	499	1	YLI1_MYCHO	HYPOTHETICAL 59.8 KD P	772	4	16.7	595	1	ZN85_HUMAN	ZINC FINGER PROTEIN 85	1.74e+03	
700	4	16.7	500	1	YK12_BACSU	HYPOTHETICAL 58.2 PROT	773	4	16.7	598	1	Y036_HUMAN	HYPOTHETICAL PROTEIN K	1.74e+03	
701	4	16.7	500	1	CN1_CAVPO	CYTOCHROME P450 XIB1 P	774	4	16.7	601	1	YAN4_SCHPO	HYPOTHETICAL 69.4 KD P	1.74e+03	
702	4	16.7	503	1	YP88_CABEL	PUTATIVE GLUCOSYLTRANS	775	4	16.7	604	1	YFIC_BACSU	HYPOTHETICAL ABC TRANS	1.74e+03	
703	4	16.7	507	1	VS15_TRYBB	VARIANT SURFACE GLYCOP	776	4	16.7	604	1	YI35_CABEL	PROBABLE G PROTEIN-COU	1.74e+03	
704	4	16.7	508	1	YB19_YEAST	PUTATIVE SERINE CARBOX	777	4	16.7	605	1	VP40_VZVD	CAPSID PROTEIN P40 (VI	1.74e+03	
705	4	16.7	509	1	YUAG_BACSU	HYPOTHETICAL 56.0 KD P	778	4	16.7	605	1	SL51_PIG	SODIUM/GLUCOSE COTRANS	1.74e+03	
706	4	16.7	513	1	Y073_TREPA	HYPOTHETICAL PROTEIN T	779	4	16.7	607	1	YSCC_YEREN	YOP PROTEINS TRANSLOCA	1.74e+03	
707	4	16.7	513	1	RP55_RHIME	RNA POLYMERASE SIGMA-5	780	4	16.7	607	1	VPP_BPH1	PROBABLE TERMINASE, AT	1.74e+03	
708	4	16.7	514	1	YLO6_CABEL	HYPOTHETICAL ALDHYDE-	781	4	16.7	609	1	HSE_SCHPO	HEAT SHOCK FACTOR PROT	1.74e+03	
709	4	16.7	515	1	YH8T_YEAST	HYPOTHETICAL 57.7 KD P	782	4	16.7	610	1	CN42_RAT	CAMP-DEPENDENT 3',5'-C	1.74e+03	
710	4	16.7	515	1	YI41_HUMAN	HYPOTHETICAL 57.7 KD P	783	4	16.7	611	1	YX45_MYCTU	HYPOTHETICAL 67.2 KD P	1.74e+03	
711	4	16.7	516	1	YP02_YEAST	HYPOTHETICAL 57.0 KD P	784	4	16.7	614	1	NTBE_RAT	SODIUM- AND CHLORIDE-D	1.74e+03	
712	4	16.7	516	1	YKCA_BACSU	HYPOTHETICAL 55.0 KD P	785	4	16.7	617	1	YACH_ECOLI	HYPOTHETICAL 69.4 KD P	1.74e+03	
713	4	16.7	517	1	Y4CC_RHISN	HYPOTHETICAL 57.8 KD P	786	4	16.7	618	1	YKRA_YEAST	HYPOTHETICAL AMINO-ACI	1.74e+03	
714	4	16.7	517	1	YAC2_YEAST	HYPOTHETICAL 58.3 KD P	787	4	16.7	621	1	Y241_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	
715	4	16.7	518	1	RTN_ECOLI	RTN PROTEIN.	788	4	16.7	623	1	YF89_YEAST	HYPOTHETICAL 70.3 KD P	1.74e+03	
716	4	16.7	520	1	YK36_YEAST	HYPOTHETICAL GTP-BINDI	789	4	16.7	623	1	Y014_MYCPN	HYPOTHETICAL ABC TRANS	1.74e+03	
717	4	16.7	523	1	YAZ4_SCHPO	HYPOTHETICAL 60.5 KD P	790	4	16.7	623	1	Y014_MYCGE	HYPOTHETICAL ABC TRANS	1.74e+03	
718	4	16.7	523	1	SYN_TREPA	ASPARAGINYL-TRNA SYNTH	791	4	16.7	623	1	Y071_MYCTU	HYPOTHETICAL 67.7 KD P	1.74e+03	
719	4	16.7	524	1	YN35_YEAST	HYPOTHETICAL 61.2 KD P	792	4	16.7	624	1	AMYG_ARXAD	GLUCOAMYLASE PRECURSOR	1.74e+03	
720	4	16.7	524	1	SAP_HUMAN	PROACTIVATOR POLYPEPTI	793	4	16.7	626	1	HTPG_BACSU	HEAT SHOCK PROTEIN HTP	1.74e+03	
721	4	16.7	524	1	FTWH_MYCTU	HYPOTHETICAL 56.3 KD F	794	4	16.7	630	1	YHJ5_YEAST	HYPOTHETICAL 72.3 KD P	1.74e+03	
722	4	16.7	525	1	YI07_METJA	HYPOTHETICAL PROTEIN M	795	4	16.7	630	1	YZ35_METJA	HYPOTHETICAL PROTEIN M	1.74e+03	
723	4	16.7	527	1	YEGH_ECOLI	HYPOTHETICAL 59.5 KD P	796	4	16.7	632	1	Y041_HUMAN	HYPOTHETICAL PROTEIN K	1.74e+03	
724	4	16.7	537	1	YR22_EBV	HYPOTHETICAL BRP2 PRO	797	4	16.7	632	1	CSX1_SCHPO	RNA-BINDING POST-TRANS	1.74e+03	
725	4	16.7	539	1	YD28_SCHPO	HYPOTHETICAL 60.7 KD P	798	4	16.7	636	1	YNR6_YEAST	HYPOTHETICAL 67.4 KD P	1.74e+03	
726	4	16.7	540	1	YK26_YEAST	HYPOTHETICAL 61.1 KD P	799	4	16.7	638	1	YQGS_BACSU	HYPOTHETICAL 73.2 KD P	1.74e+03	
727	4	16.7	544	1	YE72_METJA	HYPOTHETICAL PROTEIN M	800	4	16.7	639	1	YAJ9_SCHPO	HYPOTHETICAL 74.4 KD P	1.74e+03	
728	4	16.7	545	1	YR49_CABEL	HYPOTHETICAL 61.9 KD P	801	4	16.7	642	1	YDIF_BACSU	HYPOTHETICAL ABC TRANS	1.74e+03	
729	4	16.7	547	1	YHXB_ECOLI	HYPOTHETICAL 61.6 KD P	802	4	16.7	654	1	RPC3_YEAST	DNA-DIRECTED RNA POLYM	1.74e+03	
730	4	16.7	547	1	YQJ1_ECOLI	HYPOTHETICAL ABC TRANS	803	4	16.7	656	1	SMY1_YEAST	KINESIN-RELATED PROTEI	1.74e+03	
731	4	16.7	547	1	YKX1_CABEL	PUTATIVE SERINE/THREON	804	4	16.7	660	1	SGAT_MYCPN	SGAT PROTEIN HOMOLOG (	1.74e+03	
732	4	16.7	548	1	NRM1_BUBAR	NATURAL RESISTANCE-ASS	805	4	16.7	662	1	SL51_RABIT	SODIUM/GLUCOSE COTRANS	1.74e+03	
733	4	16.7	549	1	Y330_METJA	HYPOTHETICAL PROTEIN M	806	4	16.7	664	1	Y366_MYCPN	HYPOTHETICAL PROTEIN M	1.74e+03	
734	4	16.7	549	1	YST1_CABEL	PROBABLE 3',5'-CYCLIC	807	4	16.7	664	1	Y4FB_RHISN	HYPOTHETICAL 73.7 KD P	1.74e+03	
735	4	16.7	551	1	Y900_METJA	HYPOTHETICAL PROTEIN M	808	4	16.7	665	1	SL51_RAT	SODIUM/GLUCOSE COTRANS	1.74e+03	
736	4	16.7	552	1	YH88_YEAST	HYPOTHETICAL 63.3 KD P	809	4	16.7	666	1	YJBO_YEAST	HYPOTHETICAL 77.7 KD P	1.74e+03	
737	4	16.7	553	1	YH8A_YEAST	HYPOTHETICAL 64.4 KD P	810	4	16.7	667	1	YEO1_SCHPO	HYPOTHETICAL 76.4 KD P	1.74e+03	
738	4	16.7	554	1	CHIT_MANSE	ENDOCHITINASE PRECURSO	811	4	16.7	667	1	Y366_MYCGE	HYPOTHETICAL PROTEIN M	1.74e+03	
739	4	16.7	555	1	YFE2_YEAST	HYPOTHETICAL 62.5 KD P	812	4	16.7	669	1	YMS2_YEAST	HYPOTHETICAL 76.2 KD P	1.74e+03	
740	4	16.7	555	1	YKOC_BACSU	HYPOTHETICAL 61.5 KD P	813	4	16.7	670	1	HYFR_ECOLI	HYDROGENASE-4 TRANSERI	1.74e+03	
741	4	16.7	555	1	NRM1_CHICK	NATURAL RESISTANCE-ASS	814	4	16.7	684	1	YMW5_YEAST	HYPOTHETICAL 78.8 KD P	1.74e+03	
742	4	16.7	555	1	YQJ3_YEAST	HYPOTHETICAL 61.5 KD P	815	4	16.7	684	1	YV18_MYCTU	HYPOTHETICAL 57.3 KD P	1.74e+03	
743	4	16.7	556	1	YEAJ_ECOLI	HYPOTHETICAL 63.2 KD P	816	4	16.7	686	1	ZN07_HUMAN	ZINC FINGER PROTEIN 7	1.74e+03	
744	4	16.7	557	1	SAP_MOUSE	SULFATED GLYCOPROTEIN	817	4	16.7	689	1	ARK1_BOVIN	BETA-ADRENERGIC RECEPT	1.74e+03	
745	4	16.7	558	1	VP10_RSDV	PROTEIN S10.	818	4	16.7	691	1	YI04_YEAST	HYPOTHETICAL 80.5 KD P	1.74e+03	
746	4	16.7	558	1	YNT3_YEAST	HYPOTHETICAL 63.9 KD P	819	4	16.7	701	1	YI85_MYCGE	HYPOTHETICAL LIPOPROTE	1.74e+03	
747	4	16.7	560	1	YB65_YEAST	HYPOTHETICAL 65.9 KD P	820	4	16.7	704	1	VP51_YEAST	VACUOLAR SORTING PROTE	1.74e+03	
748	4	16.7	564	1	YK93_CABEL	HYPOTHETICAL 63.1 KD P	821	4	16.7	714	1	YHNA_YEAST	HYPOTHETICAL 80.7 KD P	1.74e+03	
749	4	16.7	567	1	YK28_METJA	HYPOTHETICAL PROTEIN M	822	4	16.7	716	1	RRP2_IATKM	RNA-DIRECTED RNA POLYM	1.74e+03	
750	4	16.7	568	1	YELG_SCHPO	PUTATIVE TRANSPORTER C	823	4	16.7	716	1	RRP2_IATKM	RNA-DIRECTED RNA POLYM	1.74e+03	
751	4	16.7	569	1	YVID_BACSU	HYPOTHETICAL 64.3 KD P	824	4	16.7	716	1	RRP2_IATKM	RNA-DIRECTED RNA POLYM	1.74e+03	
752	4	16.7	569	1	YAP3_YEAST	ASPARTIC PROTEINASE 3	825	4	16.7	716	1	RRP2_IATKM	RNA-DIRECTED RNA POLYM	1.74e+03	
753	4	16.7	569	1	YSY1_YEAST	HYPOTHETICAL 65.4 KD P	826	4	16.7	721	1	YHC6_CABEL	HYPOTHETICAL 81.4 KD P	1.74e+03	





973 4 16.7 1549 1 YJW2\_YEAST 1.74e+03  
974 4 16.7 1549 1 YIR3\_YEAST 1.74e+03  
975 4 16.7 1562 1 RPOD\_TMLVU 1.74e+03  
976 4 16.7 1615 1 RPOD\_TMLVU 1.74e+03  
977 4 16.7 1657 1 IQGA\_HUMAN 1.74e+03  
978 4 16.7 1664 1 SLF1\_CLOTH 1.74e+03  
979 4 16.7 1681 1 YE20\_YEAST 1.74e+03  
980 4 16.7 1687 1 Z142\_HUMAN 1.74e+03  
981 4 16.7 1769 1 YJW9\_YEAST 1.74e+03  
982 4 16.7 1774 1 MSAS\_PENPA 1.74e+03  
983 4 16.7 1858 1 P3K2\_DICDI 1.74e+03  
984 4 16.7 1859 1 YN79\_YEAST 1.74e+03  
985 4 16.7 1859 1 YG67\_YEAST 1.74e+03  
986 4 16.7 1868 1 YH00\_YEAST 1.74e+03  
987 4 16.7 1902 1 P2P\_LACPA 1.74e+03  
988 4 16.7 1964 1 NTC4\_MOUSE 1.74e+03  
989 4 16.7 1986 1 WALEMENI 1.74e+03  
990 4 16.7 2131 1 YCF2\_SPIOL 1.74e+03  
991 4 16.7 2198 1 YLJ2\_CAEEL 1.74e+03  
992 4 16.7 2199 1 DPOE\_SCHPO 1.74e+03  
993 4 16.7 2245 1 MYGJ\_DICDI 1.74e+03  
994 4 16.7 2280 1 YCF2\_TOBAC 1.74e+03  
995 4 16.7 2334 1 WAPA\_BACSU 1.74e+03  
996 4 16.7 2477 1 SPON\_CHICK 1.74e+03  
997 4 16.7 2733 1 RRPB\_CVMAS 1.74e+03  
998 4 16.7 3343 1 YOG7\_CAEEL 1.74e+03  
999 4 16.7 4488 1 RRAA\_CVMJH 1.74e+03  
1000 4 16.7 4644 1 DYHC\_RAT 1.74e+03

## ALIGNMENTS

RESULT 1 STANDARD: PRT; 387 AA.  
ID NK10\_HUMAN  
AC Q14943:  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (NATURAL KILLER ASSOCIATED  
DE TRANSCRIPT 10) (NKAT10).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
CC [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 96270004.  
RX DOHRING C., SAMARIDIS J., COLONNA M.;  
RT "Alternatively spliced forms of human killer inhibitory receptors.";  
RL IMMUNOGENETICS 44:227-230(1996).  
CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
CC CELL LYSIS.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC 3 C2-LIKE DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L76661; G1217709; --  
DR PFAM; PF00047; 19; 2.  
KW RECEPTOR; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW MULTIGENE FAMILY.  
FT SIGNAL 1 21 BY SIMILARITY.  
FT CHAIN 22 387 MHC CLASS I NK CELL RECEPTOR.  
FT DOMAIN 22 340 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 341 360 POTENTIAL.  
FT DOMAIN 361 387 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 42 102 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 137 202 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 237 300 IG-LIKE C2-TYPE DOMAIN.  
FT DISULFID 49 95 BY SIMILARITY.  
FT DISULFID 144 195 BY SIMILARITY.  
FT DISULFID 244 293 BY SIMILARITY.  
FT CARBOHYD 92 92 POTENTIAL.  
FT CARBOHYD 179 179 POTENTIAL.  
FT CARBOHYD 273 273 POTENTIAL.  
SQ SEQUENCE 387 AA; 42902 MW; 1412B6D7 CRC32;

Query Match 29.2%; Score 7; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 1.41e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 VTGLYK 223  
|||||||  
Qy 10 VTGLYK 16

RESULT 2 STANDARD; PRT; 211 AA.  
ID UL92\_HSV7J  
AC P52471:  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE PROTEIN U63.  
GN U63.

OS HERPES SIMPLEX VIRUS (TYPE 7 / STRAIN JI) (HSV7).  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;  
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.

RA NICHOLAS J.;  
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS TOGETHER EBV BDLF4,  
CC HVS-1 31, HSV-6 U63, AND HCMV UL92.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U43400; G113965;  
SQ SEQUENCE 211 AA; 24189 MW; 79AFDD43 CRC32;

Query Match 25.0%; Score 6; DB 1; Length 211;  
Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 144 CTFNKV 149  
|||||||  
Qy 5 CTFNKV 10

RESULT 3 STANDARD; PRT; 301 AA.  
ID NKR5\_HUMAN  
AC P43631:  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (P58 NATURAL KILLER CELL  
DE RECEPTOR CLONE CL-49) (P58 NK RECEPTOR).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE-NATURAL KILLER CELLS;  
RX MEDLINE; 95269128.

RA WAGTMANN N., BIASSONI R., CANTONI C., VERDIANI S., MALINATI M.S.,  
 RA VITALE M., BOTTINO C., MORETTA L., MORETTA A., LONG E.O.;  
 RT "Molecular clones of the p58 NK cell receptor reveal immunoglobulin-  
 RT related molecules with diversity in both the extra- and intracellular  
 RT domains";  
 RL IMMUNITY 2:439-449(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA COLONNA M., SAMARIDIS J.;  
 RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
 CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
 CC CELL LYSIS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC 2 C2-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U24079; G897911; -;  
 DR EMBL: L41347; G780312; ALT\_INIT.  
 DR PFAM: PF00047; ig; 1.  
 KW RECEPTOR: IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
 KW MULTIGENE FAMILY.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 301 MHC CLASS I NK CELL RECEPTOR.  
 FT DOMAIN 19 242 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 243 262 POTENTIAL.  
 FT DOMAIN 263 301 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 39 104 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 139 202 IG-LIKE C2-TYPE DOMAIN.  
 FT DISULFID 46 97 BY SIMILARITY.  
 FT DISULFID 146 195 BY SIMILARITY.  
 FT CARBOHYD 81 81 POTENTIAL.  
 FT CARBOHYD 175 175 POTENTIAL.  
 FT CARBOHYD 208 208 POTENTIAL.  
 SQ SEQUENCE 301 AA; 33170 MW; 86380E2E CRC32;  
 Query Match 25.0%; Score 6; DB 1; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 120 TGLYK 125  
 QY 11 TGLYK 16  
 RESULT 4  
 ID NK9\_HUMAN STANDARD; PRT; 304 AA.  
 AC Q14952; O00644;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (NATURAL KILLER ASSOCIATED  
 DE TRANSCRIPT 7) (NKAT7).  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96270004.  
 RA DOHRING C., SAMARIDIS J., COLONNA M.;  
 RT "Alternatively spliced forms of human killer inhibitory receptors.";  
 RL IMMUNOGENETICS 44:227-230(1996).  
 CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS THUS PREVENTING  
 CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
 CC CELL LYSIS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC 2 C2-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

RA BIASSONI R.;  
 RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
 CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
 CC CELL LYSIS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC 2 C2-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L76670; G1217727; -;  
 DR EMBL: X97231; E242769; -;  
 DR PFAM: PF00047; ig; 1.  
 KW RECEPTOR: IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
 KW MULTIGENE FAMILY.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 304 MHC CLASS I NK CELL RECEPTOR.  
 FT DOMAIN 22 245 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 246 264 POTENTIAL.  
 FT DOMAIN 265 304 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 42 107 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN 142 205 IG-LIKE C2-TYPE DOMAIN.  
 FT DISULFID 49 100 BY SIMILARITY.  
 FT DISULFID 149 198 BY SIMILARITY.  
 FT CARBOHYD 67 67 POTENTIAL.  
 FT CARBOHYD 84 84 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 SQ SEQUENCE 304 AA; 33717 MW; 2FD36D0D CRC32;  
 Query Match 25.0%; Score 6; DB 1; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 123 TGLYK 128  
 QY 11 TGLYK 16  
 RESULT 5  
 ID NK9\_HUMAN STANDARD; PRT; 304 AA.  
 AC Q14953;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (NATURAL KILLER ASSOCIATED  
 DE TRANSCRIPT 9) (NKAT9).  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96270004.  
 RA DOHRING C., SAMARIDIS J., COLONNA M.;  
 RT "Alternatively spliced forms of human killer inhibitory receptors.";  
 RL IMMUNOGENETICS 44:227-230(1996).  
 CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
 CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
 CC CELL LYSIS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC 2 C2-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L76672; G1217731; --  
KW RECEPTOR; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW MULTIGENE FAMILY.  
FT SIGNAL 1 21 BY SIMILARITY.  
FT CHAIN 22 304 MHC CLASS I NK CELL RECEPTOR.  
FT DOMAIN 22 245 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 246 264 POTENTIAL.  
FT DOMAIN 265 304 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 42 107 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 142 205 IG-LIKE C2-TYPE DOMAIN.  
FT DISULFID 49 100 BY SIMILARITY.  
FT DISULFID 149 198 BY SIMILARITY.  
FT CARBOHYD 67 67 POTENTIAL.  
FT CARBOHYD 84 84 POTENTIAL.  
FT CARBOHYD 178 178 POTENTIAL.  
FT CARBOHYD 223 223 POTENTIAL.  
SQ SEQUENCE 304 AA; 33644 MW; A9AE51F2 CRC32;

Query Match 25.0%; Score 6; DB 1; Length 304;

Best Local Similarity 100.0%; Pred. No. 1.16e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 123 TGLYEK 128

Qy 11 TGLYEK 16

#### RESULT 6

ID NK2\_HUMAN STANDARD; PRT; 341 AA.  
AC P43628;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (NATURAL KILLER ASSOCIATED  
DE TRANSCRIPT 2) (NKAT-2) (P58 NATURAL KILLER CELL RECEPTOR CLONE CL-6)  
DE (P58 NK RECEPTOR).  
GN NKAT2.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-NATURAL KILLER CELLS;  
RX MEDLINE; 95232526.  
RA COLONNA M., SAMARIDIS J.;  
RT "Cloning of immunoglobulin-superfamily members associated with HLA-C  
RT and HLA-B recognition by human natural killer cells.";  
RL SCIENCE 268:405-408(1995).  
RN [2]

RP SEQUENCE FROM N.A.  
RC TISSUE-NATURAL KILLER CELLS;  
RX MEDLINE; 95269128.  
RA WAGTMANN N., BIASSONI R., CANTONI C., VERDIANI S., MALNATI M.S.,  
RA VITALE M., BOTTINO C., MORETTA L., MORETTA A., LONG E.O.;  
RT "Molecular clones of the p58 NK cell receptor reveal immunoglobulin-  
RT related molecules with diversity in both the extra- and intracellular  
RT domains";  
RL IMMUNITY 2:439-449(1995).  
CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
CC CELL LYSIS.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC 2 C2-LIKE DOMAINS.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L41268; G780306; --  
KW RECEPTOR; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW MULTIGENE FAMILY.  
FT SIGNAL 1 21 BY SIMILARITY.  
FT CHAIN 22 341 MHC CLASS I NK CELL RECEPTOR.  
FT DOMAIN 22 245 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 246 265 POTENTIAL.  
FT DOMAIN 266 341 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 42 107 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 142 205 IG-LIKE C2-TYPE DOMAIN.  
FT DISULFID 49 100 BY SIMILARITY.  
FT DISULFID 149 198 BY SIMILARITY.  
FT CARBOHYD 84 84 POTENTIAL.  
FT CARBOHYD 178 178 POTENTIAL.  
FT CARBOHYD 211 211 POTENTIAL.  
SQ SEQUENCE 341 AA; 37886 MW; FACDC81C CRC32;

Query Match 25.0%; Score 6; DB 1; Length 341;

Best Local Similarity 100.0%; Pred. No. 1.16e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 123 TGLYEK 128

Qy 11 TGLYEK 16

#### RESULT 7

ID NK2\_HUMAN STANDARD; PRT; 348 AA.  
AC P43627;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (P58 NATURAL KILLER CELL  
DE RECEPTOR CLONE CL-43) (P58 NK RECEPTOR).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-NATURAL KILLER CELLS;  
RX MEDLINE; 95269128.  
RA WAGTMANN N., BIASSONI R., CANTONI C., VERDIANI S., MALNATI M.S.,  
RA VITALE M., BOTTINO C., MORETTA L., MORETTA A., LONG E.O.;  
RT "Molecular clones of the p58 NK cell receptor reveal immunoglobulin-  
RT related molecules with diversity in both the extra- and intracellular  
RT domains";  
RL IMMUNITY 2:439-449(1995).  
CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C  
CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING  
CC CELL LYSIS.  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC 2 C2-LIKE DOMAINS.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U24075; G897903; --  
KW RECEPTOR; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;  
KW MULTIGENE FAMILY.  
FT SIGNAL 1 21 BY SIMILARITY.

```
FT CHAIN 22 348 MHC CLASS I NK CELL RECEPTOR.
FT DOMAIN 22 245 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 246 264 POTENTIAL.
FT DOMAIN 246 348 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 265 348 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 42 107 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 142 205 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 49 100 BY SIMILARITY.
FT DISULFID 149 198 BY SIMILARITY.
FT CARBOHYD 84 84 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 211 211 POTENTIAL.
SQ SEQUENCE 348 AA; 38472 MW; F194AC08 CRC32;

Query Match 25.0%; Score 6; DB 1; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.16e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 123 TGLYEK 128
      |||||
QY 11 TGLYEK 16

RESULT 8
ID CAR2CANPA STANDARD; PRT; 412 AA.
AC P32950.
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CANDIDAPEPSIN 2 PRECURSOR (EC 3.4.23.24) (ASPARTATE PROTEASE 2)
DE (ACP 2).
GN SAPP2 OR ACPL
OS CANDIDA PARAPSILOSIS (YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC CANDIDACEAE; CANDIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE CHUV E18;
RX MEDLINE; 93171876.
RA DE VIRAGH P.A., SANGIARD D., TOGNI G., FALCHETTO R., MONOD M.;
RT "Cloning and sequencing of two Candida parapsilosis genes encoding
RT acid proteases.";
RL J. GEN. MICROBIOL. 139:335-342(1993).
RN [2]
RP SEQUENCE OF 62-68.
RC STRAIN-ISOLATE CHUV E18;
RX MEDLINE; 9327903.
RA FUSEK M., SMITH E.A., MONOD M., FOUNDLING S.I.;
RT "Candida parapsilosis expresses and secretes two aspartic
RT proteinases.";
RL FEBS LETT. 327:108-112(1993).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- PTM: O-GLYCOSYLATED (PROBABLE).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Z11918; G578134;
DR PIR; S20704; S20704.
DR PIR; S34581; S34581.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
DR PFAM; PF00026; asp; 1.
DR HSP; P32329; 1YPS.
KW HYDROLASE; ASPARTYL PROTEASE; GLYCOPROTEIN; ZYMOGEN; SIGNAL.
FT SIGNAL 1 216 OR 18, OR 21 (POTENTIAL).
FT PROPEP 717 61 ACTIVATION PEPTIDE.
FT CHAIN 62 412 CANDIDAPEPSIN 2.
```

```
FT ACT_SITE 93 93 BY SIMILARITY.
FT ACT_SITE 274 274 BY SIMILARITY.
FT DISULFID 108 113 BY SIMILARITY.
FT DISULFID 311 345 BY SIMILARITY.
FT CARBOHYD 53 53 POTENTIAL.
SQ SEQUENCE 412 AA; 44442 MW; 35B46196 CRC32;

Query Match 25.0%; Score 6; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 1.16e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 335 EFVLQS 340
      |||||
QY 19 EFVLQS 24

RESULT 9
ID NKR4_HUMAN STANDARD; PRT; 455 AA.
AC P43630.
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE MHC CLASS I NK CELL RECEPTOR PRECURSOR (NATURAL KILLER ASSOCIATED
DE TRANSCRIPT 4) (NKAT-4).
GN NKAT4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-NATURAL KILLER CELLS;
RX MEDLINE; 95232526.
RA COLONNA M., SAMARIDIS J.;
RT "Cloning of immunoglobulin-superfamily members associated with HLA-C
RT and HLA-B recognition by human natural killer cells.";
RL SCIENCE 268:405-408(1995).
CC -!- FUNCTION: RECEPTOR ON NATURAL KILLER (NK) CELLS FOR HLA-C
CC ALLELES. INHIBITS THE ACTIVITY OF NK CELLS THUS PREVENTING
CC CELL LYSIS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC 3 C2-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; L41270; G780310;
DR PFAM; PF00047; lg; 2.
DR RECEPTOR; IMMUNOGLOBULIN FOLD; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE;
KW MULTIGENE FAMILY.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 455 MHC CLASS I NK CELL RECEPTOR.
FT DOMAIN 22 340 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 341 360 POTENTIAL.
FT DOMAIN 361 455 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 432 102 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 137 202 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 237 300 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 49 95 BY SIMILARITY.
FT DISULFID 144 195 BY SIMILARITY.
FT DISULFID 244 293 BY SIMILARITY.
FT CARBOHYD 179 179 POTENTIAL.
FT CARBOHYD 239 239 POTENTIAL.
FT CARBOHYD 273 273 POTENTIAL.
FT CARBOHYD 306 306 POTENTIAL.
SQ SEQUENCE 455 AA; 50230 MW; BF892DD7 CRC32;

Query Match 25.0%; Score 6; DB 1; Length 455;
```



RT The coding sequences of 40 new genes (K1AA0001-K1AA0040) deduced by  
RT analysis of randomly sampled cDNA clones from human immature myeloid  
RT cell line KG-1.;  
RL DNA RES. 1:27-35(1994).  
CC -!- CATALYTIC ACTIVITY: ATP + L-LEUCINE + TRNA(LEU) = AMP +  
CC PYROPHOSPHATE + L-LEUCYL-TRNA(LEU).  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO CLASS-I AMINOCYL-TRNA SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: D21851; G434767; ALT\_INIT.  
DR EMBL: D21851; E316529; -.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; 1.  
DR PFAM: PF00133; TRNA-Syn\_1; 1.  
KW AMINOACYL-TRNA SYNTHETASE; PROTEIN BIOSYNTHESIS; LIGASE; ATP-BINDING;  
KW MITOCHONDRION; TRANSIT PEPTIDE.  
FT CHAIN 1 ? MITOCHONDRION (POTENTIAL).  
FT CHAIN 2 903 LEUCYL-TRNA SYNTHETASE.  
FT SIMILAR 92 102 "HIGH" REGION.  
FT SIMILAR 639 643 "RMSKS" REGION.  
FT BINDING 642 642 ATP (BY SIMILARITY).  
SQ SEQUENCE 903 AA; 101976 MW; 1DC1FELD CRC32;  
  
Query Match 25.0%; Score 6; DB 1; Length 903;  
Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 868 EFVLQ 873  
QY 19 EFVLQ 24  
|||||  
  
RESULT 13  
ID MP5F-CHICK STANDARD; PRT; 1450 AA.  
AC Q02173;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE M-PROTEIN, STRIATED MUSCLE.  
OS GALLUS GALLUS (CHICKEN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=EMBRYONIC PECTORALIS MUSCLE;  
RX MEDLINE; 93015907.  
RA NOGUCHI J., YANAGISAWA M., IMAMURA M., KASUYA Y., SAKURAI T.,  
RA TANAKA T., MASAKI T.;  
RT "Complete primary structure and tissue expression of chicken  
RT pectoralis M-protein.";  
RL J. BIOL. CHEM. 267:20302-20310(1992).  
CC -!- FUNCTION: IS A STRUCTURAL CONSTITUENT OF MYOFIBRILLAR M-BAND IN  
CC STRIATED MUSCLE.  
CC -!- TISSUE SPECIFICITY: IS EXPRESSED IN PECTORALIS AND CARDIAC MUSCLE.  
CC -!- DEVELOPMENTAL STAGE: CAN BE DETECTED BY DAY 10-13 IN OVO, THE  
CC CONTENT IS GRADUALLY INCREASED THROUGHOUT THE OVO DEVELOPMENT AND  
CC REACHED ITS PEAK AFTER HATCHING.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC 7 C2-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 5 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: D11474; G222833; -.  
DR PIR: A44027; A44027.  
DR PFAM: PF00041; fn3; 5.  
DR PFAM: PF00047; ig; 4.  
DR HSP: P56276; 1TK.  
KW MUSCLE PROTEIN; REPEAT; IMMUNOGLOBULIN FOLD.  
FT DOMAIN 142 238 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 266 364 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 368 464 FIBRONECTIN TYPE-III.  
FT DOMAIN 496 592 FIBRONECTIN TYPE-III.  
FT DOMAIN 597 693 FIBRONECTIN TYPE-III.  
FT DOMAIN 696 794 FIBRONECTIN TYPE-III.  
FT DOMAIN 797 898 FIBRONECTIN TYPE-III.  
FT DOMAIN 899 990 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 1002 1095 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 1126 1205 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 1225 1312 IG-LIKE C2-TYPE DOMAIN.  
FT DOMAIN 1333 1427 IG-LIKE C2-TYPE DOMAIN.  
SQ SEQUENCE 1450 AA; 163400 MW; D3AE9F0E CRC32;  
  
Query Match 25.0%; Score 6; DB 1; Length 1450;  
Best Local Similarity 100.0%; Pred. No. 1.16e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 432 VTGLYE 437  
QY 10 VTGLYE 15  
|||||  
  
RESULT 14  
ID RL11\_STRGB STANDARD; PRT; 43 AA.  
AC P52861;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE S0S RIBOSOMAL PROTEIN L11 (FRAGMENT).  
GN RPLK.  
OS STREPTOMYCES GALBUS.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;  
OC ACTINOMYCETALES; STREPTOMYCETACEAE; STREPTOMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 40480;  
RA POEHLING S., PIEPERSBERG W., WEHMEIER U.F.;  
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA  
CC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE L11P FAMILY OF RIBOSOMAL PROTEINS.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X95916; E224006; -.  
DR PROSITE: PS00359; RIBOSOMAL\_L11; PARTIAL.  
DR PFAM: PF00298; RIBOSOMAL\_L11; 1.  
KW RIBOSOMAL PROTEIN.  
FT NON\_TER 43  
SQ SEQUENCE 43 AA; 4470 MW; 868DA878 CRC32;  
  
Query Match 20.8%; Score 5; DB 1; Length 43;  
Best Local Similarity 100.0%; Pred. No. 6.13e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 7 KVTGL 11  
|||||

Qy 9 KVTGL 13

RESULT 15  
ID MONB\_DIOCU STANDARD; PRT; 50 AA.  
AC P02882:  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MONELLIN, CHAIN B (CHAIN II).  
OS DIOSCOREOPHYLLUM CUMINSII (SERENDIPITY BERRY).  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMERIOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC RANUNCULALES; RANUNCULINEAE; MENISPERMACEAE; DIOSCOREOPHYLLUM.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE; 91136778.  
RA KOHMURA M., NIO N., ARIYOSHI Y.;  
RT "Complete amino acid sequence of the sweet protein monellin.";  
RL AGRIC. BIOL. CHEM. 54:2219-2224(1990).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE; 77004858.  
RA FRANK G., ZUBER H.;  
RT "The complete amino acid sequences of both subunits of the sweet protein monellin.";  
RL HOPPE-SEYLER'S Z. PHYSIOL. CHEM. 357:585-592(1976).  
RN [3]  
RP SEQUENCE.  
RX MEDLINE; 76161292.  
RA BOHAK Z., LI S.-L.;  
RT "The structure of monellin and its relation to the sweetness of the protein.";  
RL BIOCHIM. BIOPHYS. ACTA 427:153-170(1976).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
RX MEDLINE; 87287292.  
RA OGATA C., HATADA M., TOMLINSON G., SHIN W.-C., KIM S.-H.;  
RT "Crystal structure of the intensely sweet protein monellin.";  
RL NATURE 328:739-742(1987).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).  
RX MEDLINE; 94047088.  
RA SOMOZA J.R., JIANG F., TONG L., KANG C.-H., CHO J.M., KIM S.-H.;  
RT "Two crystal structures of a potentially sweet protein. Natural monellin at 2.75-A resolution and single-chain monellin at 1.7-A resolution.";  
RL J. MOL. BIOL. 234:390-404(1993).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RA BUJACZ G., MILLER M., HARRISON R., THANKI N., GILLILAND G.L.,  
RA OGATA C., KIM S.-H., WLODAWER A.;  
RT "Structure of monellin refined to 2.3-A resolution in the orthorhombic crystal form.";  
RL ACTA CRYSTALLOGR. D 53:713-719(1997).  
RN [7]  
RP SIMILARITY TO CYSTATINS.  
RX MEDLINE; 93218009.  
RA MURZIN A.G.;  
RT "Sweet-tasting protein monellin is related to the cystatin family of thiol proteinase inhibitors.";  
RL J. MOL. BIOL. 230:689-694(1993).  
CC -!- FUNCTION: INTENSELY SWEET-TASTING PROTEIN.  
CC -!- SUBUNIT: HETERODIMER OF AN A CHAIN AND A B CHAIN.  
CC -!- SIMILARITY: DISTANTLY RELATED TO CYSTATINS.  
DR PIR: A03376; MLDIB.  
DR PIR: JH0210; JH0210.  
DR PDB: 3MON; 31-OCT-93.  
DR PDB: 4MON; 07-JUL-97.  
DR PDB: 1MOL; 31-MAY-94.  
KW SWEET-TASTE; 3D-STRUCTURE.  
FT SITE 41 41  
FT CONFLICT 49 50 BLOCKING, ABOLISHES THE SWEET TASTE.  
FT STRAND 2 2 EN -> NE (IN REF. 2).

FT STRAND 5 6  
FT TURN 8 9  
FT HELIX 10 26  
FT TURN 28 29  
FT STRAND 34 46  
SQ SEQUENCE 50 AA; 5835 MW; 0FE52EDE CRC32;

Query Match 20.8%; Score 5; DB 1; Length 50;  
Best Local Similarity 100.0%; Pred. NO. 6.13e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 TFNKV 37  
|||||  
Qy 6 TFNKV 10

Search completed: Sat Aug 28 14:42:33 1999  
Job time : 43 secs.



\*\*\*\*\*  
W A S E R E H  
\*\*\*\*\* (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:42:52 1999; MasPar time 7.37 Seconds  
Tabular output not generated.  
177.831 Million cell updates/sec

Title: >US-09-049-696-43  
Description: (1-24) from US09049696.pep  
Perfect Score: 24  
Sequence: 1 YTRCTNKVTGLYKGEFVLQS 24

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: sprembl9  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 2.545; Variance 0.362; scale 7.024

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
1	7	29.2	287	4	Q99565 NK RECEPTOR.
2	7	29.2	307	4	Q14824 KILLER CELL RECEPTOR.
3	7	29.2	335	4	Q14623 KILLER CELL RECEPTOR.
4	7	29.2	342	4	Q43803 NK RECEPTOR.
5	7	29.2	342	4	Q14622 KILLER CELL RECEPTOR.
6	7	29.2	360	4	Q99559 NK RECEPTOR.
7	7	29.2	360	4	P78401 NK-RECEPTOR.
8	7	29.2	377	4	Q14821 KILLER CELL RECEPTOR.
9	7	29.2	377	4	Q99706 NATURAL KILLER CELL RE
10	7	29.2	377	4	Q99560 NK RECEPTOR.
11	7	29.2	377	4	Q43534 NATURAL KILLER CELL IN
12	7	29.2	377	4	Q99562 NK RECEPTOR.
13	7	29.2	382	4	P78403 NK-RECEPTOR.
14	7	29.2	382	4	Q99564 NK RECEPTOR.
15	7	29.2	382	4	Q43469 NATURAL KILLER CELL RE
16	6	25.0	134	9	Q38304 L9.
17	6	25.0	160	2	Q31564 YFIV PROTEIN.
18	6	25.0	202	4	Q00382 P58 NK CELL INHIBITORY
19	6	25.0	277	5	Q44811 H04J21.2 PROTEIN.
20	6	25.0	279	10	F4i18.26 PROTEIN.

21	6	25.0	287	4	Q43756 NK RECEPTOR PRECURSOR.
22	6	25.0	287	4	Q43755 NK RECEPTOR PRECURSOR.
23	6	25.0	293	2	Q66582 HYPOTHETICAL 33.0 KD P
24	6	25.0	303	4	Q43753 NK RECEPTOR PRECURSOR.
25	6	25.0	304	4	Q14955 NK RECEPTOR FOR HLA-C
26	6	25.0	304	4	Q43754 NK RECEPTOR PRECURSOR.
27	6	25.0	327	2	Q47334 KPSF GENE.
28	6	25.0	341	4	P78402 NK-RECEPTOR.
29	6	25.0	341	4	Q14944 NKAT2A.
30	6	25.0	341	4	Q14945 NKAT2B.
31	6	25.0	341	4	Q43472 NATURAL KILLER CELL IN
32	6	25.0	348	4	Q14951 NKAT6.
33	6	25.0	390	2	Q87153 ORF22-14 PROTEIN.
34	6	25.0	390	2	Q34228 ORF43X9 PROTEIN.
35	6	25.0	436	4	Q99563 NK RECEPTOR.
36	6	25.0	436	10	B-LIKE CYCLIN.
37	6	25.0	438	7	Q95366 NK RECEPTOR.
38	6	25.0	455	7	Q95368 HLA CLASS I INHIBITORY
39	6	25.0	455	4	Q93684 NK RECEPTOR.
40	6	25.0	455	4	Q14947 NKAT4A.
41	6	25.0	455	4	Q14948 NKAT4B.
42	6	25.0	455	7	Q95367 HLA CLASS I INHIBITORY
43	6	25.0	455	4	Q13238 KIR (CL-5) NK RECEPTOR
44	6	25.0	579	5	Q60950 GOLVESIN.
45	6	25.0	26926	4	Q10466 TITIN, HEART ISOFORM N
46	5	20.8	48	2	Q86425 FERRICHRONE TRANSPORT
47	5	20.8	49	2	Q51385 HYPOTHETICAL 5.6 KD PR
48	5	20.8	86	2	P94863 RECEIVER MODULE OF PUT
49	5	20.8	88	2	Q30464 YOKE.
50	5	20.8	101	13	Q91869 CAPACITATIVE CALCULUM E
51	5	20.8	104	1	Q27485 HYPOTHETICAL 11.7 KD P
52	5	20.8	105	2	Q31868 YOTG PROTEIN.
53	5	20.8	105	9	Q64189 HYPOTHETICAL 12.5 KD P
54	5	20.8	119	2	Q32042 YRVI PROTEIN.
55	5	20.8	121	14	P89086 CAPSID (HYPERVARIABLE
56	5	20.8	131	2	Q66900 HYPOTHETICAL 15.1 KD P
57	5	20.8	133	2	Q32226 YVAD PROTEIN.
58	5	20.8	140	2	Q32448 RIBOSOMAL PROTEIN L11.
59	5	20.8	144	2	Q87733 RIBOSOMAL PROTEIN L11.
60	5	20.8	144	8	Q21032 RIBOSOMAL PROTEIN L16.
61	5	20.8	144	2	Q87085 RIBOSOMAL PROTEIN L11.
62	5	20.8	147	2	Q26086 HYPOTHETICAL 17.3 KD P
63	5	20.8	160	1	Q57903 160AA LONG HYPOTHETICA
64	5	20.8	171	1	Q26261 FERRITIN LIKE PROTEIN
65	5	20.8	172	5	Q23884 RIBOSOMAL PROTEIN L11.
66	5	20.8	173	2	Q24972 HYPOTHETICAL 20.7 KD P
67	5	20.8	180	2	Q05400 YRHH.
68	5	20.8	191	14	Q98250 MC083R.
69	5	20.8	195	2	Q83869 HYPOTHETICAL 22.1 KD P
70	5	20.8	201	2	Q33374 REGF PROTEIN.
71	5	20.8	202	4	Q00380 P50 CELL ACTIVATORY RE
72	5	20.8	202	4	Q00381 P58 NK CELL INHIBITORY
73	5	20.8	204	4	Q14950 NKAT5-DELTA-IG1.
74	5	20.8	208	5	Q18543 C39E9.2 PROTEIN.
75	5	20.8	209	14	Q58862 BEL1 PROTEIN.
76	5	20.8	209	4	Q92588 CD89_U09.
77	5	20.8	213	2	Q06937 OAC3 PROTEIN (FRAGMENT
78	5	20.8	215	2	Q54574 GLUCOSE-1-PHOSPHATE TH
79	5	20.8	219	2	Q34915 YOBE.
80	5	20.8	224	2	Q31916 YOOW PROTEIN.
81	5	20.8	224	9	Q64131 YOOW PROTEIN.
82	5	20.8	227	2	Q15150 PLATELET GLYCOPROTEIN
83	5	20.8	227	2	Q34906 YOAM.
84	5	20.8	228	2	P96657 YDDT PROTEIN.
85	5	20.8	228	2	Q31973 YOML PROTEIN.
86	5	20.8	228	9	Q64048 HYPOTHETICAL 25.4 KD P
87	5	20.8	228	2	Q34903 YKOG.
88	5	20.8	230	4	Q99561 NK RECEPTOR.
89	5	20.8	231	1	P95952 ORF C04013.
90	5	20.8	231	1	P57895 241AA LONG HYPOTHETICA
91	5	20.8	245	4	Q00547 CL-9.
92	5	20.8	249	10	Q49493 HYPOTHETICAL 27.2 KD P
93	5	20.8			

94	5	20.8	249 11	Q63489	RAP30.	8.98e+01	167	5	20.8	370 2	O66509	1108	5	20.8	249 11	Q63489	HYPOTHETICAL 42.2 KD P	8.98e+01
95	5	20.8	249 9	O80305	R PROTEIN.	8.98e+01	168	5	20.8	374 4	Q15629	1109	5	20.8	249 9	O80305	TRAM PROTEIN.	8.98e+01
96	5	20.8	253 4	Q92593	CD89 U13.	8.98e+01	169	5	20.8	377 5	O61699	1110	5	20.8	253 4	Q92593	AMPHIWITIN.	8.98e+01
97	5	20.8	255 14	O84700	29K PROTEIN.	8.98e+01	170	5	20.8	377 5	O56609	1111	5	20.8	255 14	O84700	FLAGELLIN (ORF5).	8.98e+01
98	5	20.8	257 11	O88827	PLASMA CELL MEMBRANE G	8.98e+01	171	5	20.8	381 2	O85846	1112	5	20.8	257 11	O88827	PUTATIVE AROMATIC EFFL	8.98e+01
99	5	20.8	259 1	O58747	PROBABLE NH(3)-DEPEND	8.98e+01	172	5	20.8	381 2	P76108	1113	5	20.8	259 1	O58747	FROM BASES 1506783 TO	8.98e+01
100	5	20.8	261 3	O13829	PUTATIVE SMALL NUCLEAR	8.98e+01	173	5	20.8	392 2	O86397	1114	5	20.8	261 3	O13829	PUTATIVE REGULATORY PR	8.98e+01
101	5	20.8	262 2	O55802	REGULATORY COMPONENTS	8.98e+01	174	5	20.8	393 10	O49132	1115	5	20.8	262 2	O55802	POLYPHENOL OXIDASE (FR	8.98e+01
102	5	20.8	262 2	O66813	HYPOTHETICAL 30.4 KD P	8.98e+01	175	5	20.8	396 4	O41631	1116	5	20.8	262 2	O66813	INTERLEUKIN-5 RECEPTOR	8.98e+01
103	5	20.8	263 2	O68575	PYRUVATE-FORMATE LYASE	8.98e+01	176	5	20.8	397 5	O17764	1117	5	20.8	263 2	O68575	CMGH2.2 PROTEIN.	8.98e+01
104	5	20.8	264 5	O24013	G5-LIKE ORF'S PROTEIN.	8.98e+01	177	5	20.8	397 2	O31712	1118	5	20.8	264 5	O24013	YKXZ PROTEIN.	8.98e+01
105	5	20.8	268 10	O22798	PUTATIVE CUC2 PROTEIN.	8.98e+01	178	5	20.8	398 2	O67179	1119	5	20.8	268 10	O22798	HEMOLYSIN.	8.98e+01
106	5	20.8	276 5	O52009	R186.1 PROTEIN.	8.98e+01	179	5	20.8	405 2	O66181	1120	5	20.8	276 5	O52009	TRANSPORTER.	8.98e+01
107	5	20.8	277 14	O55454	SURFACE ANTIGEN (FRAGM	8.98e+01	180	5	20.8	408 3	O67055	1121	5	20.8	277 14	O55454	CHROMOSOME XII COSMID	8.98e+01
108	5	20.8	285 2	O84834	GLUCOSE-1-PHOSPHATE TH	8.98e+01	181	5	20.8	409 11	O35220	1122	5	20.8	285 2	O84834	LYSOSOMAL SIALIDASE.	8.98e+01
109	5	20.8	285 2	O51375	CHEMOTAXIS PROTEIN MET	8.98e+01	182	5	20.8	409 11	O35657	1123	5	20.8	285 2	O51375	G9 SIALIDASE PRECURSOR	8.98e+01
110	5	20.8	287 10	O41726	VICILIN (FRAGMENT).	8.98e+01	183	5	20.8	410 14	O65352	1124	5	20.8	287 10	O41726	ORF A SPAN 522-1895.	8.98e+01
111	5	20.8	288 2	O88081	PUTATIVE GLUCOSE-1-PHO	8.98e+01	184	5	20.8	412 4	Q13558	1125	5	20.8	288 2	O88081	NR8-4AG (FRAGMENT).	8.98e+01
112	5	20.8	289 2	P97004	GLUCOSE-1-PHOSPHATE TH	8.98e+01	185	5	20.8	413 14	Q83459	1126	5	20.8	289 2	P97004	ELB PROTEIN, LARGE T A	8.98e+01
113	5	20.8	289 1	O29463	CONSERVED HYPOTHETICAL	8.98e+01	186	5	20.8	413 5	Q26438	1127	5	20.8	289 1	O29463	HEMOLIN.	8.98e+01
114	5	20.8	290 2	O05367	GLUCOSE-1-PHOSPHATE-TH	8.98e+01	187	5	20.8	414 10	Q41727	1128	5	20.8	290 2	O05367	VICILIN, 7S GLOBULIN.	8.98e+01
115	5	20.8	290 2	O65250	GLUCOSE-1-PHOSPHATE-TH	8.98e+01	188	5	20.8	415 4	O99519	1129	5	20.8	290 2	O65250	SIALIDASE PRECURSOR.	8.98e+01
116	5	20.8	291 2	O50494	RNA POLYMERASE BETA' S	8.98e+01	189	5	20.8	415 2	O31766	1130	5	20.8	291 2	O50494	YMFH PROTEIN.	8.98e+01
117	5	20.8	291 2	O68737	HYPOTHETICAL 32.3 KD P	8.98e+01	190	5	20.8	416 3	O33455	1131	5	20.8	291 2	O68737	ENDO-1,4-BETA-GLUCANAS	8.98e+01
118	5	20.8	291 4	Q92590	CD89-U02.	8.98e+01	191	5	20.8	418 5	Q22092	1132	5	20.8	291 4	Q92590	TBB-2 PROTEIN.	8.98e+01
119	5	20.8	292 2	Q54164	GLUCOSE-1-PHOSPHATE TH	8.98e+01	192	5	20.8	420 14	O64844	1133	5	20.8	292 2	Q54164	47K PROTEIN.	8.98e+01
120	5	20.8	293 2	P96448	EXPA7.	8.98e+01	193	5	20.8	420 4	Q14633	1134	5	20.8	293 2	P96448	INTERLEUKIN-5 RECEPTOR	8.98e+01
121	5	20.8	294 2	P711439	RFBA.	8.98e+01	194	5	20.8	422 4	O99684	1135	5	20.8	294 2	P711439	GROWTH FACTOR INDEPND	8.98e+01
122	5	20.8	296 14	O41275	35KDA PROTEIN.	8.98e+01	195	5	20.8	425 2	O53615	1136	5	20.8	296 14	O41275	SERINE HYDROXYMETHYLTR	8.98e+01
123	5	20.8	297 2	O69117	PUTATIVE GLUCOSE-1-PHO	8.98e+01	196	5	20.8	429 2	O59013	1137	5	20.8	297 2	O69117	HYPOTHETICAL 47.2 KD P	8.98e+01
124	5	20.8	304 4	Q14954	NK RECEPTOR FOR HLA-C	8.98e+01	197	5	20.8	439 11	O70623	1138	5	20.8	304 4	Q14954	TRANSCRIPTION FACTOR 1	8.98e+01
125	5	20.8	304 4	O43471	NATURAL KILLER CELL IN	8.98e+01	198	5	20.8	448 5	Q20028	1139	5	20.8	304 4	O43471	SIMILAR TO MAP KINASE	8.98e+01
126	5	20.8	304 7	Q30189	HLA-C SPECIFIC ACTIVAT	8.98e+01	199	5	20.8	449 2	O85377	1140	5	20.8	304 7	Q30189	ZAPD.	8.98e+01
127	5	20.8	305 10	Q93154	MYB-RELATED PROTEIN.	8.98e+01	200	5	20.8	451 13	O73719	1141	5	20.8	305 10	Q93154	T-BOA PROTEIN 5.	8.98e+01
128	5	20.8	310 3	O03193	HYPOTHETICAL 34.6 KD P	8.98e+01	201	5	20.8	452 1	O58189	1142	5	20.8	310 3	O03193	452AA LONG HYPOTHETICA	8.98e+01
129	5	20.8	312 14	O90238	REPLICATION-ASSOCIATED	8.98e+01	202	5	20.8	457 4	O60751	1143	5	20.8	312 14	O90238	SORTING NEXIN 1A.	8.98e+01
130	5	20.8	312 14	P89032	PUTATIVE PCV REPLICATI	8.98e+01	203	5	20.8	459 4	O43145	1144	5	20.8	312 14	P89032	KIAA0395 (FRAGMENT).	8.98e+01
131	5	20.8	312 14	O69217	SGG1.	8.98e+01	204	5	20.8	459 10	O80948	1145	5	20.8	312 14	O69217	PUTATIVE MYROSINASE-BI	8.98e+01
132	5	20.8	314 14	O93192	PUTATIVE REP PROTEIN (	8.98e+01	205	5	20.8	471 2	O06959	1146	5	20.8	314 14	O93192	RFBL PROTEIN.	8.98e+01
133	5	20.8	314 14	O92285	P35.8.	8.98e+01	206	5	20.8	472 10	O23226	1147	5	20.8	314 14	O92285	HYPOTHETICAL 53.5 KD P	8.98e+01
134	5	20.8	314 14	O56123	PUTATIVE REP PROTEIN.	8.98e+01	207	5	20.8	474 2	O44591	1148	5	20.8	314 14	O56123	PLASMID PTOM9 FROM ALC	8.98e+01
135	5	20.8	317 10	O23112	AP2 DOMAIN CONTAINING	8.98e+01	208	5	20.8	474 14	Q85598	1149	5	20.8	317 10	O23112	54 KDA PROTEIN.	8.98e+01
136	5	20.8	319 10	Q43792	1-AMINOCYCLOPROPANE-1-	8.98e+01	209	5	20.8	475 14	Q85159	1150	5	20.8	319 10	Q43792	PC475L.	8.98e+01
137	5	20.8	319 2	O51609	HYPOTHETICAL 37.4 KD P	8.98e+01	210	5	20.8	476 3	O59677	1151	5	20.8	319 2	O51609	HYPOTHETICAL 53.2 KD P	8.98e+01
138	5	20.8	324 2	P72671	HYPOTHETICAL 39.2 KD P	8.98e+01	211	5	20.8	496 2	O86575	1152	5	20.8	324 2	P72671	OLIGOPEPTIDE ABC TRANS	8.98e+01
139	5	20.8	324 2	O55383	2-KETOACID DEHYDROGENA	8.98e+01	212	5	20.8	497 5	O21218	1153	5	20.8	324 2	O55383	T14G10.2 (FRAGMENT).	8.98e+01
140	5	20.8	326 2	P96383	PRSA.	8.98e+01	213	5	20.8	503 2	O45834	1154	5	20.8	326 2	P96383	STARCH DEGRADING ENZYM	8.98e+01
141	5	20.8	328 2	O51149	TDP-DEOXYGLUCOSE-EPINE	8.98e+01	214	5	20.8	504 2	O87783	1155	5	20.8	328 2	O51149	TSBP PROTEIN (FRAGMENT	8.98e+01
142	5	20.8	329 10	P93732	PROLINE IMINOPEPTIDASE	8.98e+01	215	5	20.8	508 5	O45567	1156	5	20.8	329 10	P93732	F54B8.12 PROTEIN.	8.98e+01
143	5	20.8	330 2	O25902	GLYCERALDEHYDE-3-PHOSP	8.98e+01	216	5	20.8	511 2	P94618	1157	5	20.8	330 2	O25902	PUTATIVE ALPHA-AMYLASE	8.98e+01
144	5	20.8	333 4	Q15469	SOLUBLE INTERLEUKIN-5	8.98e+01	217	5	20.8	511 8	Q36791	1158	5	20.8	333 4	Q15469	CHLOROPLAST DNA FOR CP	8.98e+01
145	5	20.8	336 1	O58743	HYPOTHETICAL PROTEIN M	8.98e+01	218	5	20.8	512 14	O56677	1159	5	20.8	336 1	O58743	GLYCOPROTEIN G.	8.98e+01
146	5	20.8	336 1	O26946	GTP-BINDING PROTEIN, G	8.98e+01	219	5	20.8	517 3	O08984	1160	5	20.8	336 1	O26946	CHROMOSOME XVI READING	8.98e+01
147	5	20.8	342 6	O18744	LU-ECAM-1.	8.98e+01	220	5	20.8	518 4	O43891	1161	5	20.8	342 6	O18744	TRANSCRIPTION FACTOR T	8.98e+01
148	5	20.8	343 10	O24212	OSR40G2 GENE (FRAGMENT	8.98e+01	221	5	20.8	520 2	O82864	1162	5	20.8	343 10	O24212	AHEF.	8.98e+01
149	5	20.8	343 1	O50514	HYPOTHETICAL 37.8 KD P	8.98e+01	222	5	20.8	522 4	Q13596	1163	5	20.8	343 1	O50514	SORTING NEXIN 1.	8.98e+01
150	5	20.8	344 5	P91928	MOTOR PROTEIN (FRAGMEN	8.98e+01	223	5	20.8	522 4	O60750	1164	5	20.8	344 5	P91928	SIMILAR TO HISTIDINE A	8.98e+01
151	5	20.8	347 14	O12703	VSH-1 ASSOCIATED PROTE	8.98e+01	224	5	20.8	523 5	Q18236	1165	5	20.8	347 14	O12703	ALKYL HYDROPEROXIDE RE	8.98e+01
152	5	20.8	348 4	O43470	NATURAL KILLER CELL RE	8.98e+01	225	5	20.8	530 2	O06465	1166	5	20.8	348 4	O43470	DNA TOPOISOMERASE I (E	8.98e+01
153	5	20.8	349 4	O15301	TRANSCRIPTION FACTOR T	8.98e+01	226	5	20.8	540 2	O66893	1167	5	20.8	349 4	O15301	F411.4 PROTEIN.	8.98e+01
154	5	20.8	349 2	O05600	TRANSPOSOL TN5041 DNA.	8.98e+01	227	5	20.8	542 10	O64858	1168	5	20.8	349 2	O05600	HYPOTHETICAL 61.8 KD P	8.98e+01
155	5	20.8	356 2	O83973	GLYCEROL-3-PHOSPHATE D	8.98e+01	228	5	20.8	544 3	O14023	1169	5	20.8	356 2	O83973	FLAGELLAR FILAMENT CAP	8.98e+01
156	5	20.8	357 2	O34656	YTAAL PROTEIN.	8.98e+01	229	5	20.8	554 2	P94264	1170	5	20.8	357 2	O34656	ZK863.3 PROTEIN.	8.98e+01
157	5	20.8	358 11	O88741	GLYCOSIDE-INDUCED DI	8.98e+01	230	5	20.8	554 5	Q23651	1171	5	20.8	358 11	O88741	LPG22P.	8.98e+01
158	5	20.8	358 13	O57588	PUTATIVE TRANSCRIPTION	8.98e+01	231	5	20.8	557 3	O02908	1172	5	20.8	358 13	O57588	F411.7 PROTEIN.	8.98e+01
159	5	20.8	358 4	O75786	GANGLIOSIDE-INDUCED DI	8.98e+01	232											

240	5	20.8	631	14	P88938	ORF 50.	8.98e+01
241	5	20.8	638	2	O86147	FLGE PROTEIN (FRAGMENT	8.98e+01
242	5	20.8	647	14	Q50113	CAPSID PROTEIN VP2.	8.98e+01
243	5	20.8	663	5	O42427	FSG 120K CYS-RICH PROT	8.98e+01
244	5	20.8	665	2	O51173	FLAGELLAR HOOR-ASSOCIA	8.98e+01
245	5	20.8	674	5	P90755	C27A7.3 PROTEIN.	8.98e+01
246	5	20.8	685	14	O40935	ORF 50 (FRAGMENT).	8.98e+01
247	5	20.8	686	1	O58356	586AA LONG HYPOTHETICA	8.98e+01
248	5	20.8	691	14	O26608	HYPOTHETICAL 73.7 KD P	8.98e+01
249	5	20.8	699	3	Q12310	CHROMOSOME IV READING	8.98e+01
250	5	20.8	702	14	O36608	LARGE STRUCTURAL PROTE	8.98e+01
251	5	20.8	705	9	O64034	ABC TRANSPORTER.	8.98e+01
252	5	20.8	705	2	Q13988	SUBIANTIN 168 LANTIBIO	8.98e+01
253	5	20.8	707	10	O48701	F315.28 PROTEIN.	8.98e+01
254	5	20.8	707	2	O30671	SUBIANTIN TRANSPORTER	8.98e+01
255	5	20.8	708	14	O88442	COMPLETE GENOME.	8.98e+01
256	5	20.8	730	14	O89724	ORF 1, ORF 2 AND ORF 3	8.98e+01
257	5	20.8	758	5	O18116	SIMILAR TO HUMAN RD PR	8.98e+01
258	5	20.8	761	5	O16947	K09C6.1 PROTEIN.	8.98e+01
259	5	20.8	769	2	O68843	GLYCOSIDASE OLER.	8.98e+01
260	5	20.8	770	5	O44014	SODIUM STIBOGLUCONATE	8.98e+01
261	5	20.8	780	10	O32228	TRICHOHYALIN HOMOLOG.	8.98e+01
262	5	20.8	781	10	O80743	T13D8.9 PROTEIN.	8.98e+01
263	5	20.8	785	14	O67851	POLYMERASE.	8.98e+01
264	5	20.8	794	6	O18742	LU-ECAM-1.	8.98e+01
265	5	20.8	796	2	O25163	BIOTIN SULFOXIDE REDUC	8.98e+01
266	5	20.8	808	2	O70021	BETA-GLUCOSIDASE.	8.98e+01
267	5	20.8	815	11	P70589	C2-HC TYPE ZINC FINGER	8.98e+01
268	5	20.8	817	14	O90761	ALKALINE PHOSPHODIESTE	8.98e+01
269	5	20.8	817	13	Q07784	TITIN, MUSCLE (FRAGMEN	8.98e+01
270	5	20.8	819	2	O84211	LEUCYL TRNA SYNTHETASE	8.98e+01
271	5	20.8	820	6	O87433	LU-ECAM-1.	8.98e+01
272	5	20.8	837	2	O68454	PUTATIVE OUTER MEMBRAN	8.98e+01
273	5	20.8	853	10	O83346	OUTER MEMBRANE PROTEIN	8.98e+01
274	5	20.8	859	10	O81428	T24H24.9 PROTEIN.	8.98e+01
275	5	20.8	861	13	O23260	KINESIN-LIKE PROTEIN.	8.98e+01
276	5	20.8	864	13	Q73637	PHEROMONE RECEPTOR.	8.98e+01
277	5	20.8	869	2	P74527	AMINOPEPTIDASE.	8.98e+01
278	5	20.8	873	11	O61097	KINASE SUPPRESSOR OF R	8.98e+01
279	5	20.8	875	2	Q31978	YONG PROTEIN.	8.98e+01
280	5	20.8	875	11	O63490	RB13-6 ANTIGEN.	8.98e+01
281	5	20.8	875	11	P76765	ALKALINE PHOSPHODIESTE	8.98e+01
282	5	20.8	875	11	P76765	ALKALINE PHOSPHODIESTE	8.98e+01
283	5	20.8	875	9	O64044	YONG.	8.98e+01
284	5	20.8	875	4	O14638	PHOSPHODIESTERASE I/NU	8.98e+01
285	5	20.8	886	11	P70641	PHOSPHODIESTERASE I.	8.98e+01
286	5	20.8	886	2	O68973	PHYCOBILIPROTEIN APCE.	8.98e+01
287	5	20.8	890	5	O20404	F4P4.4 PROTEIN.	8.98e+01
288	5	20.8	905	6	O18741	LU-ECAM-1.	8.98e+01
289	5	20.8	913	11	O88826	G0B-5 PROTEIN.	8.98e+01
290	5	20.8	928	3	O74210	CHITIN SYNTHASE 2.	8.98e+01
291	5	20.8	934	1	O73954	DNA TOPOISOMERASE 1 (F	8.98e+01
292	5	20.8	949	5	P90956	T01D3.3 PROTEIN.	8.98e+01
293	5	20.8	955	2	O60681	HYPOTHETICAL 102.8 KD	8.98e+01
294	5	20.8	956	2	O66651	ISOLEUCYL-TRNA SYNTHET	8.98e+01
295	5	20.8	965	5	O15997	BMP109.	8.98e+01
296	5	20.8	994	5	O77070	SARCO/ENDOPASMIC RETI	8.98e+01
297	5	20.8	1000	6	O28737	TITIN (1000 AA) (FRAGM	8.98e+01
298	5	20.8	1023	11	P70588	C2-HC TYPE ZINC FINGER	8.98e+01
299	5	20.8	1032	10	O80491	T12M4.14 PROTEIN.	8.98e+01
300	5	20.8	1032	14	O96724	POLYPROTEIN.	8.98e+01
301	5	20.8	1042	2	O51773	ISOLEUCYL-TRNA SYNTHET	8.98e+01
302	5	20.8	1047	4	O82611	KIAA0535 PROTEIN.	8.98e+01
303	5	20.8	1064	10	O62884	T5H22.4 PROTEIN.	8.98e+01
304	5	20.8	1066	1	O58792	1066AA LONG HYPOTHETIC	8.98e+01
305	5	20.8	1073	3	Q06839	CHROMOSOME XVI COSMID	8.98e+01
306	5	20.8	1078	11	O08995	MYELIN TRANSCRIPTION F	8.98e+01
307	5	20.8	1102	10	O49552	UV-DAMAGED DNA-BINDING	8.98e+01
308	5	20.8	1106	5	O22637	T21C12.1 PROTEIN.	8.98e+01
309	5	20.8	1119	5	O24094	CARBOXYPEPTIDASE PRECU	8.98e+01
310	5	20.8	1120	10	O81509	T7M24.3 PROTEIN.	8.98e+01
311	5	20.8	1122	13	P70047	C2-HC TYPE ZINC FINGER	8.98e+01
312	5	20.8	1131	10	O23159	CU2+-TRANSPORTING ATPA	8.98e+01

VENTRICULAR ERG K+ CHA	8.98e+01
MYELIN TRANSCRIPTION F	8.98e+01
ZINC FINGER PROTEIN PN	8.98e+01
T21C9.6 PROTEIN.	8.98e+01
HYPOTHETICAL PROTEIN M	8.98e+01
T14G10.2 PROTEIN.	8.98e+01
F9D12.11 PROTEIN.	8.98e+01
(STRAIN NEW-YORK-1) P1	8.98e+01
SIMILARITY TO MULTIPLE	8.98e+01
CARBOXYPEPTIDASE SILVE	8.98e+01
COSMID 171D11.	8.98e+01
DNA POLYMERASE ALPHA 1	8.98e+01
MYELOBLAST KIAA0213 (F	8.98e+01
CONSERVED HYPOTHETICAL	8.98e+01
C25G4.10 PROTEIN.	8.98e+01
P-ATPASE.	8.98e+01
KIAA0304.	8.98e+01
MEGF6.	8.98e+01
GLUCOSYLTRANSFERASE PR	8.98e+01
180K PROTEIN.	8.98e+01
180K PROTEIN.	8.98e+01
183K PEPTIDE.	8.98e+01
EXCINUCLEASE ABC SUBUN	8.98e+01
VITELLOGENIN.	8.98e+01
TUBERIN (TUBEROUS SCLE	8.98e+01
MAL3P6.27 PROTEIN.	8.98e+01
PKS MODULE 4.	8.98e+01
'PPF3P'.	8.98e+01
TFIIIC BOX B-BINDING S	8.98e+01
HYPOTHETICAL 236.2 KD	8.98e+01
DIHYDROPYRIDINE-SENSIT	8.98e+01
DIHYDROPYRIDINE-SENSIT	8.98e+01
VOLTAGE-OPERATED CALCII	8.98e+01
INSECTICIDAL TOXIN COM	8.98e+01
FENGYCIN SYNTHETASE FE	8.98e+01
T12D8.1 PROTEIN.	8.98e+01
BETA-FILAMIN.	8.98e+01
NAD(P)H-DEPENDENT GLUT	8.98e+01
COAT PROTEIN (FRAGMENT	8.98e+01
RNA GENOME.	8.98e+01
POLYPROTEIN.	8.98e+01
REELIN.	8.98e+01
REELIN.	8.98e+01
ALPHA-AMINOADIPYL-CYST	8.98e+01
TRITHORAX PROTEIN TRX.	8.98e+01
NONCYTOPATHIC GENOMIC	8.98e+01
TITIN (FRAGMENT).	8.98e+01
EG-49E4.1 PROTEIN.	8.98e+01
UNC-89.	8.98e+01
C. ELEGANS UNC-89 (GB:	8.98e+01
TITIN (FRAGMENT).	8.98e+01
POLYKETIDE SYNTHASE.	8.98e+01
HRAR GAMMA HOMOLOG (FR	2.88e+03
RNA DEPENDENT RNA POLY	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
RNA POLYMERASE (FRAGME	2.88e+03
GLYCINE TRANSPORTER VA	2.88e+03
DOPAMINE D3 RECEPTOR I	2.88e+03
MYF5-MYF5 HOMOLOG (FRA	2.88e+03
ORF.	2.88e+03
ORF.	2.88e+03
HYPOTHETICAL 8.0 KD PR	2.88e+03
HYPOTHETICAL 8.1 KD PR	2.88e+03
HYPOTHETICAL 8.2 KD PR	2.88e+03
ORF1 PROTEIN (FRAGMENT	2.88e+03
ORF1 PROTEIN (FRAGMENT	2.88e+03

5	20.8	1144	6	O19119	
5	20.8	1182	11	O08966	
5	20.8	1188	11	Q27500	
5	20.8	1215	5	Q22649	
5	20.8	1226	1	Q58836	
5	20.8	1305	5	Q22503	
5	20.8	1322	10	O81490	
5	20.8	1362	14	O5812	
5	20.8	1372	5	P91526	
5	20.8	1406	5	Q24095	
5	20.8	1406	5	O46058	
5	20.8	1415	4	O77034	
5	20.8	1491	4	O92612	
5	20.8	1496	3	O74788	
5	20.8	1526	5	O18023	
5	20.8	1528	5	Q95022	
5	20.8	1529	4	O15022	
5	20.8	1574	11	O88281	
5	20.8	1577	2	O5265	
5	20.8	1616	14	Q98745	
5	20.8	1616	14	Q93058	
5	20.8	1616	14	Q98748	
5	20.8	1786	2	O84337	
5	20.8	1807	5	O17428	
5	20.8	1814	11	O61037	
5	20.8	1864	5	O77366	
5	20.8	1937	2	O30482	
5	20.8	1953	3	O13450	
5	20.8	2109	4	Q12789	
5	20.8	2172	5	O09515	
5	20.8	2270	4	O14581	
5	20.8	2272	11	Q61290	
5	20.8	2312	4	O15878	
5	20.8	2504	2	O85160	
5	20.8	2554	2	O30981	
5	20.8	2561	5	O46025	
5	20.8	2602	4	O75369	
5	20.8	3097	5	O61143	
5	20.8	3140	14	Q84925	
5	20.8	3141	14	P89038	
5	20.8	3199	14	O85074	
5	20.8	3460	4	P78509	
5	20.8	3461	11	Q60841	
5	20.8	3722	2	P94873	
5	20.8	3726	5	Q27255	
5	20.8	3898	14	O11994	
5	20.8	4650	4	O15598	
5	20.8	5327	5	O76891	
5	20.8	6632	5	Q17362	
5	20.8	6642	5	O01761	
5	20.8	6875	6	O28733	
5	20.8	10223	2	Q54296	
4	16.7	34	11	O64344	
4	16.7	36	14	O90722	
4	16.7	39	14	Q83756	
4	16.7	39	14	Q83747	
4	16.7	39	14	Q83746	
4	16.7	39	14	Q83751	
4	16.7	39	14	Q83753	
4	16.7	39	14	Q83754	
4	16.7	39	14	Q83752	
4	16.7	39	14	Q83755	
4	16.7	55	14	Q86777	
4	16.7	62	11	Q63322	
4	16.7	63	11	O63875	
4	16.7	64	13	Q91959	
4	16.7	64	14	P87727	
4	16.7	64	14	Q67554	
4	16.7	70	14	O90377	
4	16.7	71	14	O90375	
4	16.7	72	14	O90376	
4	16.7	74	14	O91627	
4	16.7	74	14	O91611	

386	4	16.7	74 14	091608	ORF1 PROTEIN (FRAGMENT	2.88e+03	459	177 14	090422	COAT PROTEIN (FRAGMENT	2.88e+03
387	4	16.7	74 14	091607	ORF1 PROTEIN (FRAGMENT	2.88e+03	460	177 14	090281	COAT PROTEIN (FRAGMENT	2.88e+03
388	4	16.7	74 14	091599	ORF1 (FRAGMENT)	2.88e+03	461	177 14	090184	COAT PROTEIN (FRAGMENT	2.88e+03
389	4	16.7	74 14	070861	ORF1 (FRAGMENT)	2.88e+03	462	177 14	090171	COAT PROTEIN (FRAGMENT	2.88e+03
390	4	16.7	80 1	059660	ALAO H+ ATPASE, SUBUNI	2.88e+03	463	177 14	089916	COAT PROTEIN (FRAGMENT	2.88e+03
391	4	16.7	83 2	074271	HYPOTHETICAL 10.1 KD P	2.88e+03	464	177 14	090045	CORE COAT PROTEIN (FRA	2.88e+03
392	4	16.7	86 14	090998	GP120 (FRAGMENT)	2.88e+03	465	177 14	089914	COAT PROTEIN (FRAGMENT	2.88e+03
393	4	16.7	88 14	088296	RNA POLYMERASE (FRAGME	2.88e+03	466	177 14	090183	COAT PROTEIN (FRAGMENT	2.88e+03
394	4	16.7	88 14	088298	RNA POLYMERASE (FRAGME	2.88e+03	467	177 14	090181	COAT PROTEIN (FRAGMENT	2.88e+03
395	4	16.7	88 14	088293	RNA POLYMERASE (FRAGME	2.88e+03	468	177 14	090180	COAT PROTEIN (FRAGMENT	2.88e+03
396	4	16.7	88 14	088297	RNA POLYMERASE (FRAGME	2.88e+03	469	177 14	090046	CORE COAT PROTEIN (FRA	2.88e+03
397	4	16.7	88 14	088300	RNA POLYMERASE (FRAGME	2.88e+03	470	177 14	090182	COAT PROTEIN (FRAGMENT	2.88e+03
398	4	16.7	88 14	087258	HYPOTHETICAL 10.2 KD P	2.88e+03	471	177 14	072678	CORE COAT PROTEIN (FRA	2.88e+03
399	4	16.7	97 14	092618	RNA POLYMERASE (FRAGME	2.88e+03	472	177 14	072353	COAT PROTEIN (FRAGMENT	2.88e+03
400	4	16.7	99 14	041009	NORWALK CALCIVIRUS GEN	2.88e+03	473	177 14	072680	CORE COAT PROTEIN (FRA	2.88e+03
401	4	16.7	103 2	072389	HISTIDINE KINASE HOMOL	2.88e+03	474	177 14	072683	CORE COAT PROTEIN (FRA	2.88e+03
402	4	16.7	104 14	092286	P12.1.	2.88e+03	475	180 2	025828	HYPOTHETICAL 21.0 KD P	2.88e+03
403	4	16.7	104 14	091922	ENVELOPE GLYCOPROTEIN	2.88e+03	476	181 14	085120	VIRAL PROTEIN 1 (FRAGM	2.88e+03
404	4	16.7	104 14	056124	STRAIN PWMS PCV, COMPL	2.88e+03	477	183 8	036333	PMV-ORF183.	2.88e+03
405	4	16.7	105 14	067021	DI-3 PROTEIN.	2.88e+03	478	186 2	071436	RFBD.	2.88e+03
406	4	16.7	105 14	092467	ACMNPV ORF108.	2.88e+03	479	195 14	082682	MATRIX PROTEIN.	2.88e+03
407	4	16.7	111 8	003470	CYTOCHROME B (FRAGMENT	2.88e+03	480	196 2	048831	PLASMOD P1 GENES REPA	2.88e+03
408	4	16.7	114 5	016864	SUPEROXIDE DISMUTASE (	2.88e+03	481	200 4	078513	CYSTEINE DIOXYGENASE.	2.88e+03
409	4	16.7	116 2	053750	COP PROTEIN.	2.88e+03	482	201 2	086026	PUTATIVE PROPIONYL-COA	2.88e+03
410	4	16.7	118 2	032175	YUSI PROTEIN.	2.88e+03	483	207 5	091283	SIMILAR TO ATPASE DELT	2.88e+03
411	4	16.7	119 14	091601	ORF1 (FRAGMENT).	2.88e+03	484	214 14	009513	DNA POLYMERASE (EC 2.7	2.88e+03
412	4	16.7	119 14	091602	ORF1 (FRAGMENT).	2.88e+03	485	215 12	097975	70-KDA HEAT SHOCK PROT	2.88e+03
413	4	16.7	121 14	091950	LONG ORF (FRAGMENT).	2.88e+03	486	218 14	055589	NSP5.	2.88e+03
414	4	16.7	122 14	091948	LONG ORF (FRAGMENT).	2.88e+03	487	218 11	061330	TRANSIENTLY-EXPRESSED	2.88e+03
415	4	16.7	123 14	091854	ORF-3.	2.88e+03	488	218 2	025829	DNA POLYMERASE III DEL	2.88e+03
416	4	16.7	123 14	081877	COMPLETE GENOME SEQUEN	2.88e+03	489	222 14	073128	ENVELOPE GLYCOPROTEIN	2.88e+03
417	4	16.7	123 14	090299	STRUCTURAL PROTEIN.	2.88e+03	490	224 14	081995	CAPSID PROTEIN L1 (FRA	2.88e+03
418	4	16.7	126 10	042078	PROTEIN KINASE RECEPTO	2.88e+03	491	225 14	081991	CAPSID PROTEIN L1 (FRA	2.88e+03
419	4	16.7	126 14	091595	AC3 PROTEIN	2.88e+03	492	230 14	089748	NONSTRUCTURAL PROTEIN	2.88e+03
420	4	16.7	131 5	062058	C9orf9.1 PROTEIN.	2.88e+03	493	230 14	058264	NS1 NONSTRUCTURAL PROT	2.88e+03
421	4	16.7	133 14	090243	GP48 (FRAGMENT).	2.88e+03	494	230 14	089286	NONSTRUCTURAL PROTEIN	2.88e+03
422	4	16.7	133 4	099777	HYPOTHETICAL 14.1 KD P	2.88e+03	495	230 14	089284	NONSTRUCTURAL PROTEIN	2.88e+03
423	4	16.7	135 3	043016	GTP-BINDING PROTEIN 1	2.88e+03	496	233 14	090305	ENVELOPE PROTEIN.	2.88e+03
424	4	16.7	136 2	031949	YONI PROTEIN.	2.88e+03	497	234 3	060148	HYPOTHETICAL 26.2 KD P	2.88e+03
425	4	16.7	138 5	025905	MALARIA ANTIGEN (FRAGM	2.88e+03	498	236 14	067552	ORF.	2.88e+03
426	4	16.7	140 14	091994	POLYPROTEIN (FRAGMENT)	2.88e+03	499	238 3	074551	HYPOTHETICAL 26.9 KD P	2.88e+03
427	4	16.7	140 14	091996	POLYPROTEIN (FRAGMENT)	2.88e+03	500	239 2	053678	ENTEROTOXIN (FRAGMENT)	2.88e+03
428	4	16.7	140 14	086737	TAX PROTEIN (FRAGMENT)	2.88e+03	501	239 5	025772	ASPARAGINE-RICH ANTIGE	2.88e+03
429	4	16.7	141 14	067037	HEMAGGLUTININ (FRAGMEN	2.88e+03	502	239 2	006534	ENTEROTOXIN TYPE C (SE	2.88e+03
430	4	16.7	141 14	067036	HEMAGGLUTININ (FRAGMEN	2.88e+03	503	239 2	006533	ENTEROTOXIN TYPE C (SE	2.88e+03
431	4	16.7	141 14	067038	HEMAGGLUTININ (FRAGMEN	2.88e+03	504	240 2	032200	VVQV PROTEIN.	2.88e+03
432	4	16.7	141 9	080187	PUTATIVE HOLIN.	2.88e+03	505	246 10	038831	AUXIN-INDUCIBLE IAA13.	2.88e+03
433	4	16.7	142 14	081306	POLYPROTEIN PRECURSOR	2.88e+03	506	247 2	066254	ABC TRANSPORTER.	2.88e+03
434	4	16.7	142 14	066301	RNA POLYMERASE (FRAGME	2.88e+03	507	251 14	089123	COAT PROTEIN.	2.88e+03
435	4	16.7	142 14	066302	RNA POLYMERASE (FRAGME	2.88e+03	508	251 1	028912	PHOSPHATE ABC TRANSPOR	2.88e+03
436	4	16.7	142 2	056497	ALPHA SUBUNIT OF DINIT	2.88e+03	509	251 14	091197	AV1 PROTEIN.	2.88e+03
437	4	16.7	144 10	038784	MANNOSE-BINDING LECTIN	2.88e+03	510	257 14	009351	COAT PROTEIN.	2.88e+03
438	4	16.7	147 14	087630	41KBP-FRAGMENT FROM LE	2.88e+03	511	257 14	009350	COAT PROTEIN.	2.88e+03
439	4	16.7	148 14	065601	REVERSE TRANSCRIPTASE	2.88e+03	512	259 5	062050	COBF11.4 PROTEIN.	2.88e+03
440	4	16.7	149 14	012671	ONCOPROTEIN.	2.88e+03	513	263 8	020674	CYTOCHROME B (FRAGMENT	2.88e+03
441	4	16.7	149 14	090310	COAT PROTEIN CORE REGI	2.88e+03	514	263 8	020674	CYTOCHROME B (FRAGMENT	2.88e+03
442	4	16.7	153 14	089331	LECTIN HOMOLOG.	2.88e+03	515	263 11	062798	PAIRED-LIKE HOMEODOMA	2.88e+03
443	4	16.7	156 14	090309	COAT PROTEIN CORE REGI	2.88e+03	516	269 5	061890	K10C9.7 PROTEIN.	2.88e+03
444	4	16.7	159 14	090308	COAT PROTEIN CORE REGI	2.88e+03	517	271 2	073788	PHOSPHATE TRANSPORT AT	2.88e+03
445	4	16.7	160 3	007921	CHROMOSOME XII READING	2.88e+03	518	272 2	030546	ACCG.	2.88e+03
446	4	16.7	164 14	093037	L1 CAPSID PROTEIN (FRA	2.88e+03	519	275 1	029611	AGMATINASE (SPEB).	2.88e+03
447	4	16.7	164 14	091195	L1 CAPSID PROTEIN (FRA	2.88e+03	520	277 5	094949	YOLK PROTEIN 1 (FRAGME	2.88e+03
448	4	16.7	166 5	024943	CATHEPSIN L-LIKE PROTE	2.88e+03	521	278 14	088166	COAT PROTEIN (FRAGMENT	2.88e+03
449	4	16.7	167 1	058924	HYPOTHETICAL PROTEIN M	2.88e+03	522	278 2	051225	UTP--GLUCOSE-1-PHOSPHA	2.88e+03
450	4	16.7	169 10	004922	GLUTATHIONE PEROXIDASE	2.88e+03	523	281 2	054259	GLA 0 GENES STRB1, STR	2.88e+03
451	4	16.7	170 14	067942	DNA POLYMERASE (EC 2.7	2.88e+03	524	282 14	039238	THYMIDYLATE SYNTHASE H	2.88e+03
452	4	16.7	172 5	016014	IMMUNOGLOBULIN G BINDI	2.88e+03	525	285 14	012931	UL34 HOMOLOG (FRAGMENT	2.88e+03
453	4	16.7	173 2	030426	F47B8.1 PROTEIN.	2.88e+03	526	285 14	089432	(HHV-6).	2.88e+03
454	4	16.7	173 2	031042	D10L DEHYDRATASE. SMALL	2.88e+03	527	287 14	090758	SNAP.	2.88e+03
455	4	16.7	177 14	090280	COAT PROTEIN (FRAGMENT	2.88e+03	528	287 11	035278	ADP-RIBOSYLTRANSFERASE	2.88e+03
456	4	16.7	177 14	090307	CORE COAT PROTEIN CORE REGI	2.88e+03	529	291 14	067929	DNA POLYMERASE (EC 2.7	2.88e+03
457	4	16.7	177 14	090308	CORE COAT PROTEIN (FRA	2.88e+03	530	291 14	067932	DNA POLYMERASE (EC 2.7	2.88e+03
458	4	16.7	177 14	090282	COAT PROTEIN (FRAGMENT	2.88e+03	531	295 2	054097	HYPOTHETICAL 32.1 KD P	2.88e+03

532	4	16.7	296 14	Q86358	G PROTEIN (FRAGMENT).	2.88e+03	605	4	16.7	372 1	028929	CITRATE SYNTHASE (CITZ	2.88e+03
533	4	16.7	296 2	P74223	GLUTAMINE-BINDING PROT	2.88e+03	606	4	16.7	372 14	Q84527	SIMILAR TO HUMAN ORNIT	2.88e+03
534	4	16.7	298 14	Q82062	GLYCOPROTEIN.	2.88e+03	607	4	16.7	373 14	Q90697	GLYCOPROTEIN E (FRAGME	2.88e+03
535	4	16.7	298 14	Q82060	GLYCOPROTEIN.	2.88e+03	608	4	16.7	374 8	O20607	CYTCHROME B (FRAGMENT	2.88e+03
536	4	16.7	298 14	Q82061	(MAD-1-93).	2.88e+03	609	4	16.7	374 2	O83496	HYPOTHETICAL 42.1 K D	2.88e+03
537	4	16.7	298 14	O82061	ATTACHMENT PROTEIN.	2.88e+03	610	4	16.7	379 8	O20616	CYTCHROME B.	2.88e+03
538	4	16.7	299 2	O50492	GLUTAMATE PERMEASE.	2.88e+03	611	4	16.7	379 8	O20611	CYTCHROME B.	2.88e+03
539	4	16.7	303 10	Q69028	P GENE PRODUCT.	2.88e+03	612	4	16.7	381 5	Q21831	FIBRONECTIN ATTACHMENT	2.88e+03
540	4	16.7	303 10	Q38789	MANNOSE-BINDING LECTIN	2.88e+03	613	4	16.7	381 5	Q21831	FIBRONECTIN ATTACHMENT	2.88e+03
541	4	16.7	303 2	Q54009	SPAO.	2.88e+03	614	4	16.7	383 14	Q67012	HEMAGGLUTININ (FRAGMEN	2.88e+03
542	4	16.7	305 14	Q67856	DNA POLYMERASE (EC 2.7	2.88e+03	615	4	16.7	385 1	O27038	POLYPROTEIN PRECURSOR	2.88e+03
543	4	16.7	305 14	Q67859	DNA POLYMERASE (EC 2.7	2.88e+03	616	4	16.7	385 1	O27038	ATP SYNTHASE, SUBUNIT	2.88e+03
544	4	16.7	309 10	P93160	URICASE (EC 1.7.3.3).	2.88e+03	617	4	16.7	386 14	O92096	NS5B PROTEIN (FRAGMENT	2.88e+03
545	4	16.7	312 2	Q43382	YVOD.	2.88e+03	618	4	16.7	387 14	O92463	ACMNPV ORF103.	2.88e+03
546	4	16.7	313 4	Q76000	DNA POLYMERASE (EC 2.7	2.88e+03	619	4	16.7	387 14	O92463	NUCLEOPROTEIN.	2.88e+03
547	4	16.7	315 8	O03345	DYCHROMIR.1 (OLFACTORY R	2.88e+03	620	4	16.7	391 14	Q84649	NUCLEOCAPSID PROTEIN (	2.88e+03
548	4	16.7	319 5	O45990	CYTCHROME B. (FRAGMENT	2.88e+03	621	4	16.7	395 10	O23674	CHALCONE SYNTHASE HOMO	2.88e+03
549	4	16.7	320 2	P77987	ZK1037.8 PROTEIN.	2.88e+03	622	4	16.7	395 10	O23674	POLYPROTEIN PRECURSOR	2.88e+03
550	4	16.7	320 2	O35008	DNA POLYMERASE FAMILY	2.88e+03	623	4	16.7	397 14	O86351	INNER CAPSID PROTEIN V	2.88e+03
551	4	16.7	325 14	Q83964	YTOA.	2.88e+03	624	4	16.7	397 14	O86341	INNER CAPSID PROTEIN V	2.88e+03
552	4	16.7	326 14	Q67009	HAI CHAIN.	2.88e+03	625	4	16.7	397 14	O86187	VP6.	2.88e+03
553	4	16.7	326 14	Q67009	POLYPROTEIN PRECURSOR	2.88e+03	626	4	16.7	397 14	O86187	INNER CAPSID PROTEIN V	2.88e+03
554	4	16.7	326 5	O24944	CATHEPSIN L-LIKE PROTE	2.88e+03	627	4	16.7	397 14	O86187	POLYKETIDE SYNTHASE.	2.88e+03
555	4	16.7	327 14	O82491	HAEMAGGLUTININ (FRAGME	2.88e+03	628	4	16.7	398 2	O55224	POLYKETIDE SYNTHASE.	2.88e+03
556	4	16.7	327 14	O82492	HAEMAGGLUTININ (FRAGME	2.88e+03	629	4	16.7	401 4	O13147	ABL INTERACTOR 2.	2.88e+03
557	4	16.7	330 10	Q42975	BETA GLUCOSIDASE (EC 3	2.88e+03	630	4	16.7	402 2	O53371	TRANSPOSASE.	2.88e+03
558	4	16.7	331 13	O93538	LIPOPROTEIN SIRA.	2.88e+03	631	4	16.7	410 10	O80355	SLR1 (FRAGMENT).	2.88e+03
559	4	16.7	332 10	O93538	LACTATE DEHYDROGENASE-	2.88e+03	632	4	16.7	411 14	O92626	GLYCOPROTEIN G.	2.88e+03
560	4	16.7	332 10	O89099	PYRUVATE FORMATE-LYASE	2.88e+03	633	4	16.7	412 10	O34057	ORF26.	2.88e+03
561	4	16.7	334 14	O36257	XYLOGLUCAN ENDOTRANSG	2.88e+03	634	4	16.7	415 3	O74183	ACYL- (ACYL CARRIER PRO	2.88e+03
562	4	16.7	335 14	O91205	TAX PROTEIN (FRAGMENT)	2.88e+03	635	4	16.7	416 13	Q90419	FLAVOHENOGLOBIN.	2.88e+03
563	4	16.7	336 1	O28939	HAEMAGGLUTININ (FRAGME	2.88e+03	636	4	16.7	417 5	O23226	TIGGT-WINKLE HEDGEHOG.	2.88e+03
564	4	16.7	337 14	P90463	PYRUVATE FORMATE-LYASE	2.88e+03	637	4	16.7	421 3	O74995	W09C2.4 PROTEIN.	2.88e+03
565	4	16.7	340 11	O54733	ORF 70.	2.88e+03	638	4	16.7	423 11	P70338	TFIIB SUBUNIT P47.	2.88e+03
566	4	16.7	340 11	O54733	HAEMAGGLUTININ (FRAGMEN	2.88e+03	639	4	16.7	424 14	P87626	GROWTH FACTOR INDEPND	2.88e+03
567	4	16.7	344 14	O84075	CATHEPSIN S PRECURSOR	2.88e+03	640	4	16.7	427 5	O77101	41KB FRAGMENT FROM LE	2.88e+03
568	4	16.7	344 14	O84074	INFLUENZA A/TAIWAN/1/8	2.88e+03	641	4	16.7	432 14	O90311	NUCLEAR TRANSCRIPTION	2.88e+03
569	4	16.7	344 14	O67137	HEMAGGLUTININ (FRAGMEN	2.88e+03	642	4	16.7	439 2	O85833	MAJOR CAPSID PROTEIN V	2.88e+03
570	4	16.7	344 14	O90752	HEMAGGLUTININ (FRAGMEN	2.88e+03	643	4	16.7	441 10	O22282	HYPOTHETICAL 48.2 KD P	2.88e+03
571	4	16.7	344 14	O12293	SERINE PROTEASE INHIBI	2.88e+03	644	4	16.7	443 11	O35942	T517.1 PROTEIN.	2.88e+03
572	4	16.7	344 14	O67064	HAEMAGGLUTININ HAI SUBU	2.88e+03	645	4	16.7	445 2	O69240	NIMA-RELATED KINASE 2.	2.88e+03
573	4	16.7	344 14	O67067	HAEMAGGLUTININ (FRAGME	2.88e+03	646	4	16.7	445 5	O23091	NRD GENE.	2.88e+03
574	4	16.7	344 14	O67096	HAEMAGGLUTININ (FRAGME	2.88e+03	647	4	16.7	446 14	O92101	SIMILAR TO NADPH DEHYD	2.88e+03
575	4	16.7	344 14	O67114	HAEMAGGLUTININ PRECURS	2.88e+03	648	4	16.7	448 14	O90421	NS5B PROTEIN (FRAGMENT	2.88e+03
576	4	16.7	344 14	O07779	HAEMAGGLUTININ PRECURS	2.88e+03	649	4	16.7	449 14	O65877	GLYCOPROTEIN G PRECURS	2.88e+03
577	4	16.7	344 14	O67065	HAEMAGGLUTININ PRECURS	2.88e+03	650	4	16.7	451 10	O04309	ORF5 (FRAGMENT).	2.88e+03
578	4	16.7	344 14	O67058	HAEMAGGLUTININ (FRAGME	2.88e+03	651	4	16.7	452 10	O48715	JASMONATE INDUCIBLE PR	2.88e+03
579	4	16.7	344 14	O07768	HAEMAGGLUTININ (FRAGME	2.88e+03	652	4	16.7	454 14	O92098	PUTATIVE FLAVONOL 3-O-	2.88e+03
580	4	16.7	344 14	O07753	HAEMAGGLUTININ PRECURS	2.88e+03	653	4	16.7	457 5	O60994	NS5B PROTEIN (FRAGMENT	2.88e+03
581	4	16.7	344 14	O67124	HAEMAGGLUTININ PRECURS	2.88e+03	654	4	16.7	462 14	P89917	CHITINASE.	2.88e+03
582	4	16.7	344 14	O67121	HAEMAGGLUTININ PRECURS	2.88e+03	655	4	16.7	465 11	O88354	C-TERMINAL PORTION OF	2.88e+03
583	4	16.7	344 14	O07919	HAEMAGGLUTININ PRECURS	2.88e+03	656	4	16.7	470 10	O65028	HEART PANCREATIC LIPAS	2.88e+03
584	4	16.7	344 14	O67054	HAEMAGGLUTININ PRECURS	2.88e+03	657	4	16.7	472 13	O93342	POLLEN-SPECIFIC 1-AMIN	2.88e+03
585	4	16.7	344 6	O77833	CHEMOKINE RECEPTOR CCR	2.88e+03	658	4	16.7	472 2	O06039	MATRIX METALLOPROTEINA	2.88e+03
586	4	16.7	344 2	O68270	LIPOPOLYSACCHARIDE COR	2.88e+03	659	4	16.7	475 14	O84501	EPSK.	2.88e+03
587	4	16.7	347 2	P97115	ACETOHYDROXYACID ISOME	2.88e+03	660	4	16.7	475 1	O26479	SIMILAR TO E. COLI YHE	2.88e+03
588	4	16.7	350 14	O12292	HAEMAGGLUTININ HAI SUBU	2.88e+03	661	4	16.7	477 13	O93322	O-ANTIGEN TRANSPORTER	2.88e+03
589	4	16.7	352 6	O18771	CCR5 RECEPTOR (FRAGMEN	2.88e+03	662	4	16.7	478 14	O89242	ETS-RELATED TRANSCRIPT	2.88e+03
590	4	16.7	352 6	O18770	CCR5 RECEPTOR (FRAGMEN	2.88e+03	663	4	16.7	480 10	O24609	GLYCOPROTEIN C.	2.88e+03
591	4	16.7	353 14	O12294	HAEMAGGLUTININ HAI SUBU	2.88e+03	664	4	16.7	480 4	O73472	1-AMINOCYCLOPROPANE-1-	2.88e+03
592	4	16.7	354 14	O87759	TRANSMEMBRANE GLYCOPRO	2.88e+03	665	4	16.7	482 10	O65328	TUMOROUS IMAGINAL DISC	2.88e+03
593	4	16.7	355 1	O27789	HYPOTHETICAL 40.4 KD P	2.88e+03	666	4	16.7	482 2	O45452	ACC SYNTHASE.	2.88e+03
594	4	16.7	358 14	P89324	TRANSMEMBRANE GLYCOPRO	2.88e+03	667	4	16.7	483 14	O92102	MOB PROTEIN.	2.88e+03
595	4	16.7	360 10	O65039	CYSTINE ENDOPEPTIDASE	2.88e+03	668	4	16.7	485 2	O34233	NS5B PROTEIN (FRAGMENT	2.88e+03
596	4	16.7	361 14	P89127	ACL PROTEIN.	2.88e+03	669	4	16.7	487 14	O92960	ORF56X5 PROTEIN.	2.88e+03
597	4	16.7	362 14	O92406	P43=ACMNPV ORF39.	2.88e+03	670	4	16.7	487 14	O92960	GP160 (FRAGMENT).	2.88e+03
598	4	16.7	362 14	O91309	CYTOSINE METHYLTRANSFER	2.88e+03	671	4	16.7	488 14	O92100	NS5B PROTEIN (FRAGMENT	2.88e+03
599	4	16.7	364 14	O90338	MATRIX PROTEIN.	2.88e+03	672	4	16.7	488 2	O85775	PUTATIVE REGULATORY PR	2.88e+03
600	4	16.7	364 14	O83839	MATRIX PROTEIN.	2.88e+03	673	4	16.7	490 14	O92427	LEF-9=ACMNPV ORF62.	2.88e+03
601	4	16.7	364 14	O92330	MATRIX PROTEIN.	2.88e+03	674	4	16.7	491 5	O25288	MAJOR SURFACE GLYCOPRO	2.88e+03
602	4	16.7	364 14	O83832	MATRIX PROTEIN.	2.88e+03	675	4	16.7	496 14	O98457	GENOME, PARTIAL SEQUEN	2.88e+03
603	4	16.7	364 14	O83835	MATRIX PROTEIN.	2.88e+03	676	4	16.7	497 5	O18193	CATALASE (EC 1.11.1.6)	2.88e+03
604	4	16.7	371 14	O86843	FIBER.	2.88e+03	677	4	16.7	500 2	O52526	POLYSACCHARIDE EXPORT	2.88e+03
										501 3	Q07470	CHROMOSOME IV READING	2.88e+03

678	4	16.7	506	2	O07855	SQUALENE SYNTHASE.	2.88e+03	751	4	16.7	741	14	Q02358	PROBABLE COAT PROTEIN	2.88e+03
679	4	16.7	507	14	O89985	TRANSMEMBRANE GLYCOPRO	2.88e+03	752	4	16.7	743	14	O09517	DNA POLYMERASE (EC 2.7	2.88e+03
680	4	16.7	507	14	O89983	TRANSMEMBRANE GLYCOPRO	2.88e+03	753	4	16.7	743	5	O77281	EG-EG0002.1 PROTEIN.	2.88e+03
681	4	16.7	507	14	O89984	TRANSMEMBRANE GLYCOPRO	2.88e+03	754	4	16.7	746	5	O23056	SIMILAR TO THE CARNITI	2.88e+03
682	4	16.7	507	14	O89126	LI PROTEIN.	2.88e+03	755	4	16.7	747	3	Q04377	D9719.5P.	2.88e+03
683	4	16.7	510	1	O52629	BETA-GALACTOSIDASE.	2.88e+03	756	4	16.7	749	13	O93598	TRANSCRIPTION FACTOR.	2.88e+03
684	4	16.7	513	14	O29099	NS5B PROTEIN (FRAGMENT	2.88e+03	757	4	16.7	752	11	P70627	NAAG-PEPTIDASE (GLUTAM	2.88e+03
685	4	16.7	513	14	O81220	STRUCTURAL PROTEIN (FR	2.88e+03	758	4	16.7	753	5	O17590	C16C2.3 PROTEIN.	2.88e+03
686	4	16.7	516	14	O88804	MAJOR CAPSID PROTEIN L	2.88e+03	759	4	16.7	754	5	P91063	COSMID C17H11.	2.88e+03
687	4	16.7	517	14	O83378	POLYMERASE AND ENVELOP	2.88e+03	760	4	16.7	761	5	O18458	SIMILAR TO DROSOPHILA	2.88e+03
688	4	16.7	521	5	O77213	PUTATIVE HISTONE DEACE	2.88e+03	761	4	16.7	770	14	O90364	ORF1 PROTEIN.	2.88e+03
689	4	16.7	521	14	P87597	41KBP FRAGMENT FROM LE	2.88e+03	762	4	16.7	776	14	O91565	POL PROTEIN.	2.88e+03
690	4	16.7	522	14	O91334	GLYCOPROTEIN E.	2.88e+03	763	4	16.7	781	14	Q90200	VP1 STRUCTURAL PROTEIN	2.88e+03
691	4	16.7	522	4	O60240	PERILIPIN.	2.88e+03	764	4	16.7	781	14	Q90200	VP1 AND VP2 STRUCTURAL	2.88e+03
692	4	16.7	523	3	O07904	CHROMOSOME XII READING	2.88e+03	765	4	16.7	782	5	O26104	RHOPTRY ASSOCIATED PRO	2.88e+03
693	4	16.7	524	14	O36367	MINOR CAPSID SCAFFOLD	2.88e+03	766	4	16.7	785	14	O93119	PUTATIVE 90.4K PROTEIN	2.88e+03
694	4	16.7	524	2	O86638	HYPOTHETICAL 56.3 KD P	2.88e+03	767	4	16.7	786	2	O31874	YOSO PROTEIN.	2.88e+03
695	4	16.7	527	5	O93762	F53C11.4 PROTEIN.	2.88e+03	768	4	16.7	787	14	O09511	DNA POLYMERASE (EC 2.7	2.88e+03
696	4	16.7	533	11	O08569	ACROSOMAL MATRIX COMPO	2.88e+03	769	4	16.7	791	14	Q98261	MC094R.	2.88e+03
697	4	16.7	537	2	P71879	HYPOTHETICAL 57.1 KD P	2.88e+03	770	4	16.7	796	13	Q42590	MIS5P.	2.88e+03
698	4	16.7	538	5	O77100	NUCLEAR TRANSCRIPTION	2.88e+03	771	4	16.7	801	14	O09504	DNA POLYMERASE (EC 2.7	2.88e+03
699	4	16.7	545	14	O90054	ORF2.	2.88e+03	772	4	16.7	801	14	O09505	DNA POLYMERASE (EC 2.7	2.88e+03
700	4	16.7	549	2	P71651	HYPOTHETICAL 60.3 KD P	2.88e+03	773	4	16.7	803	5	O17493	B0393.4 PROTEIN.	2.88e+03
701	4	16.7	560	14	O06347	REVERSE TRANSCRIPTASE	2.88e+03	774	4	16.7	805	5	Q18628	C44F1.5 PROTEIN (FRAGM	2.88e+03
702	4	16.7	564	2	P94746	YHCK PROTEIN.	2.88e+03	775	4	16.7	806	14	O91522	POL PROTEIN.	2.88e+03
703	4	16.7	565	14	O67010	POLYPROTEIN PRECURSOR.	2.88e+03	776	4	16.7	811	14	O91582	POL PROTEIN.	2.88e+03
704	4	16.7	566	14	O67007	POLYPROTEIN PRECURSOR.	2.88e+03	777	4	16.7	825	14	O67885	DNA POLYMERASE (EC 2.7	2.88e+03
705	4	16.7	566	14	O80894	HEMAGGLUTININ.	2.88e+03	778	4	16.7	827	14	O91539	POL PROTEIN.	2.88e+03
706	4	16.7	566	14	O98095	HEMAGGLUTININ.	2.88e+03	779	4	16.7	827	14	O91514	POL PROTEIN.	2.88e+03
707	4	16.7	566	14	O98092	HEMAGGLUTININ.	2.88e+03	780	4	16.7	827	14	O91587	POL PROTEIN.	2.88e+03
708	4	16.7	566	14	O67040	HEMAGGLUTININ.	2.88e+03	781	4	16.7	828	14	O91545	POL PROTEIN.	2.88e+03
709	4	16.7	566	14	O67039	HEMAGGLUTININ.	2.88e+03	782	4	16.7	828	14	O91569	POL PROTEIN.	2.88e+03
710	4	16.7	566	14	P88836	HAEMAGGLUTININ HA (FRA	2.88e+03	783	4	16.7	828	14	O56554	DNA POLYMERASE (EC 2.7	2.88e+03
711	4	16.7	566	14	Q82774	HAEMAGGLUTININ PRECURS	2.88e+03	784	4	16.7	830	14	O91574	POL PROTEIN.	2.88e+03
712	4	16.7	566	14	Q82503	HEMAGGLUTININ.	2.88e+03	785	4	16.7	832	14	Q67919	DNA POLYMERASE (EC 2.7	2.88e+03
713	4	16.7	566	14	O82834	HEMAGGLUTININ PRECURSO	2.88e+03	786	4	16.7	832	14	P87744	DNA POLYMERASE (EC 2.7	2.88e+03
714	4	16.7	567	10	O29045	POSSIBLE ALDEHYDE DECA	2.88e+03	787	4	16.7	832	14	Q67913	DNA POLYMERASE (EC 2.7	2.88e+03
715	4	16.7	567	10	O23116	ETHYLENE-INSENSITIVE3-	2.88e+03	788	4	16.7	832	14	O96846	COMPLETE GENOME.	2.88e+03
716	4	16.7	568	11	O54902	NATURAL RESISTANCE ASS	2.88e+03	789	4	16.7	832	14	O67907	DNA POLYMERASE (EC 2.7	2.88e+03
717	4	16.7	568	14	O84419	PBCV-1 HYALURONIC ACID	2.88e+03	790	4	16.7	832	14	Q67882	DNA POLYMERASE (EC 2.7	2.88e+03
718	4	16.7	569	14	Q27546	ENVELOPE GLYCOPROTEIN	2.88e+03	791	4	16.7	832	14	Q81169	DNA POLYMERASE (EC 2.7	2.88e+03
719	4	16.7	572	4	O16797	NADP-DEPENDENT MALIC E	2.88e+03	792	4	16.7	832	14	Q67892	DNA POLYMERASE (EC 2.7	2.88e+03
720	4	16.7	575	14	O39506	GLYCOPROTEIN E.	2.88e+03	793	4	16.7	832	14	Q67874	DNA POLYMERASE (EC 2.7	2.88e+03
721	4	16.7	580	1	O06634	NETRIN-2 LIKE PROTEIN.	2.88e+03	794	4	16.7	832	14	O67948	DNA POLYMERASE (EC 2.7	2.88e+03
722	4	16.7	585	1	O06504	V-ATPASE A SUBUNIT (EC	2.88e+03	795	4	16.7	832	14	O11885	DNA POLYMERASE (EC 2.7	2.88e+03
723	4	16.7	586	3	O60136	HYPOTHETICAL 66.1 KD P	2.88e+03	796	4	16.7	832	14	O56655	DNA POLYMERASE (EC 2.7	2.88e+03
724	4	16.7	586	5	O01895	CODED FOR BY C. ELEGAN	2.88e+03	797	4	16.7	833	14	O91585	POL PROTEIN.	2.88e+03
725	4	16.7	593	14	O84248	T-ANTIGEN.	2.88e+03	798	4	16.7	838	14	Q81165	DNA POLYMERASE (EC 2.7	2.88e+03
726	4	16.7	597	2	O51037	PHOSPHOGLUCOMUTASE (FE	2.88e+03	799	4	16.7	838	14	P88802	DNA POLYMERASE (EC 2.7	2.88e+03
727	4	16.7	599	5	O25276	GLYCOPROTEIN 63 PRECUR	2.88e+03	800	4	16.7	839	14	O92446	P95-ACMNPV ORF83.	2.88e+03
728	4	16.7	607	5	O45936	Y43F4B.7 PROTEIN.	2.88e+03	801	4	16.7	842	14	O69590	DNA POLYMERASE (EC 2.7	2.88e+03
729	4	16.7	609	5	O22533	HYPOTHETICAL 71.0 KD P	2.88e+03	802	4	16.7	842	14	O69602	DNA POLYMERASE (EC 2.7	2.88e+03
730	4	16.7	611	6	O62763	ACETYLCHOLINESTERASE C	2.88e+03	803	4	16.7	843	14	O39882	DNA POLYMERASE (EC 2.7	2.88e+03
731	4	16.7	614	14	P87611	41KBP FRAGMENT FROM LE	2.88e+03	804	4	16.7	843	14	Q67937	DNA POLYMERASE (EC 2.7	2.88e+03
732	4	16.7	619	14	O90051	LARGE T ANTIGEN (FRAGM	2.88e+03	805	4	16.7	843	14	O39644	DNA POLYMERASE (EC 2.7	2.88e+03
733	4	16.7	625	10	O22681	GL1 HOMOLOG.	2.88e+03	806	4	16.7	843	14	O09509	DNA POLYMERASE (EC 2.7	2.88e+03
734	4	16.7	632	2	O65748	NADH DEHYDROGENASE I C	2.88e+03	807	4	16.7	843	14	O42041	DNA POLYMERASE (EC 2.7	2.88e+03
735	4	16.7	633	14	O92055	NS5B PROTEIN (FRAGMENT	2.88e+03	808	4	16.7	843	14	O91539	POL PROTEIN.	2.88e+03
736	4	16.7	634	3	O74879	CONSERVED HYPOTHETICAL	2.88e+03	809	4	16.7	843	14	O91524	POL PROTEIN.	2.88e+03
737	4	16.7	638	10	O39953	POLY(A)-MRNA BINDING P	2.88e+03	810	4	16.7	843	14	O91527	POL PROTEIN.	2.88e+03
738	4	16.7	642	1	O38928	HYPOTHETICAL PROTEIN M	2.88e+03	811	4	16.7	843	14	O81137	DNA POLYMERASE (EC 2.7	2.88e+03
739	4	16.7	643	14	O83249	PHOSPHOTRANSFERASE (FR	2.88e+03	812	4	16.7	843	14	O81131	DNA POLYMERASE (EC 2.7	2.88e+03
740	4	16.7	650	10	O65465	PROTEIN KINASE - LIKE	2.88e+03	813	4	16.7	843	14	O91589	POL PROTEIN.	2.88e+03
741	4	16.7	666	2	P72531	PENICILLIN-BINDING PRO	2.88e+03	814	4	16.7	843	14	Q67925	DNA POLYMERASE (EC 2.7	2.88e+03
742	4	16.7	666	2	P95804	PENICILLIN-BINDING PROT	2.88e+03	815	4	16.7	843	14	O91518	POL PROTEIN.	2.88e+03
743	4	16.7	686	4	O35195	VILLIN-LIKE PROTEIN.	2.88e+03	816	4	16.7	843	14	O91511	POL PROTEIN.	2.88e+03
744	4	16.7	690	3	O74874	PUTATIVE TRANSCRIPTION	2.88e+03	817	4	16.7	843	14	O81124	DNA POLYMERASE (EC 2.7	2.88e+03
745	4	16.7	691	11	O35804	JANUS PROTEIN TYROSINE	2.88e+03	818	4	16.7	843	14	O91516	POL PROTEIN.	2.88e+03
746	4	16.7	698	14	O67889	DNA POLYMERASE (EC 2.7	2.88e+03	819	4	16.7	843	14	O81120	DNA POLYMERASE (EC 2.7	2.88e+03
747	4	16.7	710	13	O91402	HEPATOCYTE GROWTH FACT	2.88e+03	820	4	16.7	843	14	O81137	DNA POLYMERASE (EC 2.7	2.88e+03
748	4	16.7	716	14	O91742	PA POLYMERASE SUBUNIT.	2.88e+03	821	4	16.7	843	14	O81107	DNA POLYMERASE (EC 2.7	2.88e+03
749	4	16.7	719	14	P87541	VIRAL POLYPROTEIN PREC	2.88e+03	822	4	16.7	843	14	O91551	POL PROTEIN.	2.88e+03
750	4	16.7	720	14	O92094	NS5B PROTEIN (FRAGMENT	2.88e+03	823	4	16.7	843	14	O91547	POL PROTEIN.	2.88e+03

824	4	16.7	843 14	081141	DNA POLYMERASE (EC 2.7	2.88e+03	897	4	16.7	1041 14	Q02442	GENOME POLYPROTEIN [CO	2.88e+03
825	4	16.7	843 14	091549	POL PROTEIN.	2.88e+03	898	4	16.7	1054 14	Q65810	P125 PROTEIN (FRAGMENT	2.88e+03
826	4	16.7	843 14	091541	POL PROTEIN.	2.88e+03	899	4	16.7	1055 14	Q89928	REVERSE TRANSCRIPTASE.	2.88e+03
827	4	16.7	843 14	039877	DNA POLYMERASE (EC 2.7	2.88e+03	900	4	16.7	1057 14	Q98204	MC036R.	2.88e+03
828	4	16.7	843 14	091572	POL PROTEIN.	2.88e+03	901	4	16.7	1059 14	Q76630	POL POLYPROTEIN (FRAGM	2.88e+03
829	4	16.7	843 14	081111	DNA POLYMERASE (EC 2.7	2.88e+03	902	4	16.7	1060 14	Q88016	POL POLYPROTEIN.	2.88e+03
830	4	16.7	843 14	091584	POL PROTEIN.	2.88e+03	903	4	16.7	1061 4	Q13420	STEROID RECEPTOR COACT	2.88e+03
831	4	16.7	843 14	091576	POL PROTEIN.	2.88e+03	904	4	16.7	1067 14	P85904	FROM AFRICAN GREEN MON	2.88e+03
832	4	16.7	843 14	091578	POL PROTEIN.	2.88e+03	905	4	16.7	1082 5	Q77431	EG:118B3.2 PROTEIN.	2.88e+03
833	4	16.7	843 14	091580	POL PROTEIN.	2.88e+03	906	4	16.7	1103 14	Q41928	SSDNA BINDING PROTEIN.	2.88e+03
834	4	16.7	843 14	Q69594	DNA POLYMERASE (EC 2.7	2.88e+03	907	4	16.7	1138 14	Q65716	ALPHA PROTEIN.	2.88e+03
835	4	16.7	843 14	Q39571	DNA POLYMERASE (EC 2.7	2.88e+03	908	4	16.7	1142 4	Q14324	FAST MYBP-C.	2.88e+03
836	4	16.7	843 14	Q03766	DNA POLYMERASE (EC 2.7	2.88e+03	909	4	16.7	1148 5	Q77203	CYTOSOLIC REGULATOR PI	2.88e+03
837	4	16.7	843 14	Q39666	DNA POLYMERASE (EC 2.7	2.88e+03	910	4	16.7	1153 5	Q93598	F29D11.2 PROTEIN.	2.88e+03
838	4	16.7	843 14	Q69616	DNA POLYMERASE (EC 2.7	2.88e+03	911	4	16.7	1156 14	Q93209	POL PROTEIN.	2.88e+03
839	4	16.7	843 14	Q81099	DNA POLYMERASE (EC 2.7	2.88e+03	912	4	16.7	1160 14	Q86928	ORF1 (FRAGMENT).	2.88e+03
840	4	16.7	843 14	Q81134	DNA POLYMERASE (EC 2.7	2.88e+03	913	4	16.7	1187 13	Q93284	POL POLYPROTEIN (FRAGM	2.88e+03
841	4	16.7	843 14	Q81116	DNA POLYMERASE (EC 2.7	2.88e+03	914	4	16.7	1198 3	Q13353	CLASS IV CHITIN SYNTHA	2.88e+03
842	4	16.7	843 14	Q91563	POL PROTEIN.	2.88e+03	915	4	16.7	1201 14	Q39510	VIRAL POLYPROTEIN PREC	2.88e+03
843	4	16.7	843 14	Q91562	POL PROTEIN.	2.88e+03	916	4	16.7	1211 3	Q14072	PROBABLE CALCIUM-TRANS	2.88e+03
844	4	16.7	843 14	Q91561	POL PROTEIN.	2.88e+03	917	4	16.7	1220 14	Q41894	POL POLYPROTEIN.	2.88e+03
845	4	16.7	843 14	Q69605	DNA POLYMERASE (EC 2.7	2.88e+03	918	4	16.7	1227 4	P78504	JAGGED 1 (TRANSMEMBRAN	2.88e+03
846	4	16.7	843 14	Q91558	POL PROTEIN.	2.88e+03	919	4	16.7	1247 14	Q90371	STRUCTURAL POLYPROTEIN	2.88e+03
847	4	16.7	843 14	Q91559	POL PROTEIN.	2.88e+03	920	4	16.7	1249 5	Q27534	W09C2.3 PROTEIN.	2.88e+03
848	4	16.7	843 14	Q91554	POL PROTEIN.	2.88e+03	921	4	16.7	1251 13	Q91365	QN1 ORF (FRAGMENT).	2.88e+03
849	4	16.7	843 14	Q91556	POL PROTEIN.	2.88e+03	922	4	16.7	1254 14	Q88613	STRUCTURAL POLYPROTEIN	2.88e+03
850	4	16.7	845 14	Q91533	POL PROTEIN.	2.88e+03	923	4	16.7	1277 14	Q08547	TRANS-ACTING TRANSCRIP	2.88e+03
851	4	16.7	845 14	Q91536	POL PROTEIN.	2.88e+03	924	4	16.7	1289 14	Q98246	MC079R.	2.88e+03
852	4	16.7	845 14	Q67895	DNA POLYMERASE (EC 2.7	2.88e+03	925	4	16.7	1302 14	Q65734	CORE PROTEIN.	2.88e+03
853	4	16.7	845 14	Q81096	DNA POLYMERASE (EC 2.7	2.88e+03	926	4	16.7	1302 14	Q65741	RNA-DIRECTED RNA POLYM	2.88e+03
854	4	16.7	850 14	Q97009	ENVELOPE GLYCOPROTEIN	2.88e+03	927	4	16.7	1302 14	Q65740	RNA-DIRECTED RNA POLYM	2.88e+03
855	4	16.7	850 14	Q97008	ENVELOPE GLYCOPROTEIN	2.88e+03	928	4	16.7	1302 14	Q65739	RNA-DIRECTED RNA POLYM	2.88e+03
856	4	16.7	851 14	Q70145	ENVELOPE GLYCOPROTEIN	2.88e+03	929	4	16.7	1317 11	Q35211	APOPTOSIS-ASSOCIATED T	2.88e+03
857	4	16.7	852 14	Q90103	ENVELOPE GLYCOPROTEIN	2.88e+03	930	4	16.7	1339 5	Q26048	BETA-1,3-D-GLUCAN BIND	2.88e+03
858	4	16.7	852 14	Q41546	ENVELOPE GLYCOPROTEIN	2.88e+03	931	4	16.7	1340 14	Q88274	POLYPROTEIN (FRAGMENT)	2.88e+03
859	4	16.7	853 14	Q70202	ENVELOPE GLYCOPROTEIN	2.88e+03	932	4	16.7	1361 14	Q83331	S GLYCOPROTEIN.	2.88e+03
860	4	16.7	853 14	Q92594	RAS GTPASE-ACTIVATING	2.88e+03	933	4	16.7	1374 10	Q22148	PUTATIVE REVERSE TRANS	2.88e+03
861	4	16.7	854 14	Q36362	GLYCOPROTEIN B.	2.88e+03	934	4	16.7	1386 5	Q77327	MALP3.10 PROTEIN.	2.88e+03
862	4	16.7	855 14	Q69660	ENVELOPE GLYCOPROTEIN	2.88e+03	935	4	16.7	1404 5	Q62170	F14F4.3 PROTEIN.	2.88e+03
863	4	16.7	856 14	P88523	ENVELOPE GLYCOPROTEIN	2.88e+03	936	4	16.7	1404 4	Q13177	E1F4I.1 (EUKARYOTIC INI	2.88e+03
864	4	16.7	858 2	Q25722	CELL DIVISION PROTEIN	2.88e+03	937	4	16.7	1411 2	P76701	RHSB CORE PROTEIN WITH	2.88e+03
865	4	16.7	856 2	Q84500	DNA POLYMERASE I.	2.88e+03	938	4	16.7	1432 14	Q28751	GAG-POL PRECURSOR.	2.88e+03
866	4	16.7	873 14	Q68287	NS3 POLYPROTEIN (FRAGM	2.88e+03	939	4	16.7	1442 14	Q83858	RNA POLYMERASE.	2.88e+03
867	4	16.7	877 14	Q68288	NS3 POLYPROTEIN (FRAGM	2.88e+03	940	4	16.7	1453 14	Q91603	SPIKE PROTEIN.	2.88e+03
868	4	16.7	884 4	Q00302	ERPROT 213-21.	2.88e+03	941	4	16.7	1507 5	Q24298	DE-CADHERIN.	2.88e+03
869	4	16.7	887 4	Q16507	COLLAGEN VII (FRAGMENT	2.88e+03	942	4	16.7	1520 4	Q15087	KIAA0383 (FRAGMENT).	2.88e+03
870	4	16.7	893 14	Q69062	HERPESVIRUS TYPE 6 (HH	2.88e+03	943	4	16.7	1542 4	Q15035	KIAA0321 (FRAGMENT).	2.88e+03
871	4	16.7	896 14	Q96547	POLYPROTEIN OF DEFECTI	2.88e+03	944	4	16.7	1550 4	Q92547	MVELOBLAST KIAA0259 (F	2.88e+03
872	4	16.7	899 14	Q83393	GAG PRO POL POLYPROTEI	2.88e+03	945	4	16.7	1616 14	Q90211	180K PROTEIN.	2.88e+03
873	4	16.7	901 14	Q69465	GLYCOPROTEIN B PRECURS	2.88e+03	946	4	16.7	1642 4	Q14869	YOTTAO.	2.88e+03
874	4	16.7	902 14	Q69095	GLYCOPROTEIN B PRECURS	2.88e+03	947	4	16.7	1648 14	Q86366	186K PROTEIN.	2.88e+03
875	4	16.7	904 14	Q90734	POLYPROTEIN.	2.88e+03	948	4	16.7	1675 14	Q86366	194K POLYPEPTIDE.	2.88e+03
876	4	16.7	904 14	Q69526	GLYCOPROTEIN B.	2.88e+03	949	4	16.7	1675 14	Q90301	POLYPROTEIN.	2.88e+03
877	4	16.7	912 14	Q90278	ENVELOPE PROTEIN.	2.88e+03	950	4	16.7	1690 3	Q74835	PUTATIVE RNA BIOGENES	2.88e+03
878	4	16.7	934 14	Q91858	HEXON PROTEIN.	2.88e+03	951	4	16.7	1693 14	Q81876	COMPLETE GENOME SEQUEN	2.88e+03
879	4	16.7	937 14	Q83128	HEXON GENE.	2.88e+03	952	4	16.7	1693 14	Q39221	NONSTRUCTURAL POLYPROT	2.88e+03
880	4	16.7	943 2	Q52149	ORF 1.	2.88e+03	953	4	16.7	1695 5	Q62604	VOLTAGE-GATED SODIUM C	2.88e+03
881	4	16.7	950 2	Q50470	ACONITATE HYDRATASE.	2.88e+03	954	4	16.7	1704 14	Q67724	RNA-DEPENDENT RNA POLY	2.88e+03
882	4	16.7	950 2	Q50470	MC090R. (FRAGMENT).	2.88e+03	955	4	16.7	1763 14	Q66913	NON-STRUCTURAL PROTEIN	2.88e+03
883	4	16.7	950 14	Q98257	SIMILAR TO HHV8A U89 A	2.88e+03	956	4	16.7	1792 5	Q18983	D2045.2 PROTEIN.	2.88e+03
884	4	16.7	958 14	Q57151	POLYPROTEIN (FRAGMENT)	2.88e+03	957	4	16.7	1810 13	Q90824	CYTOTACTIN 200KD PRECU	2.88e+03
885	4	16.7	964 14	Q65757	BZIP HISTIDINE KINASE.	2.88e+03	958	4	16.7	1869 5	Q15763	HYBRID HISTIDINE KINAS	2.88e+03
886	4	16.7	978 2	Q07831	ENDO-N-ACETYLNEURAMINI	2.88e+03	959	4	16.7	2004 14	P87509	RNA-DEPENDENT RNA POLY	2.88e+03
887	4	16.7	984 9	Q80102	PUTATIVE MINOR STRUCTU	2.88e+03	960	4	16.7	2014 5	Q01703	T25C12.3 PROTEIN.	2.88e+03
888	4	16.7	1000 9	Q80181	GLUTAMATE RECEPTOR CHA	2.88e+03	961	4	16.7	2146 3	Q59897	POLYKETIDE SYNTHASE.	2.88e+03
889	4	16.7	1007 11	Q61625	AH ANTIGEN (FRAGMENT).	2.88e+03	962	4	16.7	2162 14	Q91940	RNA DEPENDENT RNA POLY	2.88e+03
890	4	16.7	1017 4	Q13171	POL POLYPROTEIN (FRAGM	2.88e+03	963	4	16.7	2258 14	P90245	POLYPROTEIN PRECURSOR.	2.88e+03
891	4	16.7	1019 14	P89154	120 KD ANTIGEN.	2.88e+03	964	4	16.7	2258 14	Q93128	RNAL POLYPROTEIN.	2.88e+03
892	4	16.7	1022 2	Q52658	REVERSE TRANSCRIPTASE.	2.88e+03	965	4	16.7	2334 14	Q96725	RNA.	2.88e+03
893	4	16.7	1022 14	Q87965	REVERSE TRANSCRIPTASE.	2.88e+03	966	4	16.7	2343 6	Q62730	FACTOR VIII.	2.88e+03
894	4	16.7	1022 14	Q87956	DNA POLYMERASE.	2.88e+03	967	4	16.7	2493 14	Q90163	NONSTRUCTURAL POLYPROT	2.88e+03
895	4	16.7	1023 14	Q92787	PHOSPHOLIPASE D1.	2.88e+03	968	4	16.7	2493 14	Q65592	PUTATIVE NONSTRUCTURAL	2.88e+03
896	4	16.7	1036 11	Q08684		2.88e+03	969	4	16.7	2498 14	Q85924	POLYPROTEIN 1.	2.88e+03

```
970      4 16.7 2543 2 031784      2.88e+03
971      4 16.7 2635 14 040942      2.88e+03
972      4 16.7 2643 5 001552      2.88e+03
973      4 16.7 2733 14 039226      2.88e+03
974      4 16.7 2870 14 P89968      2.88e+03
975      4 16.7 2925 14 089163      2.88e+03
976      4 16.7 3005 14 056074      2.88e+03
977      4 16.7 3010 14 069829      2.88e+03
978      4 16.7 3010 14 069949      2.88e+03
979      4 16.7 3011 14 003463      2.88e+03
980      4 16.7 3014 14 091936      2.88e+03
981      4 16.7 3019 14 092529      2.88e+03
982      4 16.7 3037 14 068749      2.88e+03
983      4 16.7 3070 14 089906      2.88e+03
984      4 16.7 3222 14 085071      2.88e+03
985      4 16.7 3222 14 085057      2.88e+03
986      4 16.7 3410 14 089277      2.88e+03
987      4 16.7 3410 14 089278      2.88e+03
988      4 16.7 3411 14 091857      2.88e+03
989      4 16.7 3493 14 039734      2.88e+03
990      4 16.7 3898 14 092365      2.88e+03
991      4 16.7 3898 14 092366      2.88e+03
992      4 16.7 3898 14 068871      2.88e+03
993      4 16.7 3898 14 092872      2.88e+03
994      4 16.7 3898 14 092364      2.88e+03
995      4 16.7 3898 14 068964      2.88e+03
996      4 16.7 3898 14 068965      2.88e+03
997      4 16.7 3898 14 068966      2.88e+03
998      4 16.7 4096 4 013327      2.88e+03
999      4 16.7 4930 2 031783      2.88e+03
1000     4 16.7 4981 5 077372      2.88e+03
```

## ALIGNMENTS

```
RESULT 1
ID Q99565; PRELIMINARY; PRT; 287 AA.
AC Q99565;
DT 01-MAY-1997 (TREMBREL. 03, CREATED)
DT 01-MAY-1997 (TREMBREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE NK RECEPTOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOID;
RA BIASSONI R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X97232; E246050; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 287 AA; 31737 MW; CB497599 CRC32;
```

Query Match 29.2%; Score 7; DB 4; Length 287;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 122 VTGLYEK 128
      |||||
Qy 10 VTGLYEK 16
```

```
RESULT 2
ID O14624; PRELIMINARY; PRT; 307 AA.
AC O14624;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE KILLER CELL RECEPTOR.
GN KIR103.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
```

```
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A., STEFFENS U., PALANISAMY N., CHAGANTI R.S.K.,
RA DUPONT B.;
RL TISSUE ANTIGENS 0:0-0(1997).
DR EMBL; AF002982; G2443484; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 307 AA; 33631 MW; EC40A7DA CRC32;
```

Query Match 29.2%; Score 7; DB 4; Length 307;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 119 VTGLYEK 125
      |||||
Qy 10 VTGLYEK 16
```

```
RESULT 3
ID O14623; PRELIMINARY; PRT; 325 AA.
AC O14623;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE KILLER CELL RECEPTOR.
GN KIR103.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A., STEFFENS U., PALANISAMY N., CHAGANTI R.S.K.,
RA DUPONT B.;
RL TISSUE ANTIGENS 0:0-0(1997).
DR EMBL; AF002981; G2443482; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 325 AA; 35609 MW; 2C0FA13B CRC32;
```

Query Match 29.2%; Score 7; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 119 VTGLYEK 125
      |||||
Qy 10 VTGLYEK 16
```

```
RESULT 4
ID O43803; PRELIMINARY; PRT; 342 AA.
AC O43803;
DT 01-JUN-1998 (TREMBREL. 06, CREATED)
DT 01-JUN-1998 (TREMBREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)
DE NK RECEPTOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOID;
RA BIASSONI R.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y13054; E322404; -.
SQ SEQUENCE 342 AA; 37305 MW; 1DD5856A CRC32;
```

Query Match 29.2%; Score 7; DB 4; Length 342;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 119 VTGLYEK 125
      |||||
Qy 10 VTGLYEK 16
```



```

RA SELVAKUMAR A.;
RT "Polymorphism and domain variability of human killer cell inhibitory
RL IMMUNOL. REV. 155:183-196(1997).
DR EMBL; U7394; G1890666; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 360 AA; 39740 MW; B9C85111 CRC32;

Query Match      29.2%; Score 7; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.69e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYK 125
QY 10 VTGLYK 16
|||||

RESULT 8
ID O14621 PRELIMINARY; PRT; 377 AA.
AC O14821;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KILLER CELL RECEPTOR.
GN KIR103
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A., STEFFENS U., PALANISAMY N., CHAGANTI R.S.K.,
RA DUPONT B.;
RL TISSUE ANTIGENS 0:0-0(1997).
DR EMBL; AF002979; G2443478; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 377 AA; 41523 MW; C0651356 CRC32;

Query Match      29.2%; Score 7; DB 4; Length 377;
Best Local Similarity 100.0%; Pred. No. 1.69e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYK 125
QY 10 VTGLYK 16
|||||

RESULT 9
ID Q9706 PRELIMINARY; PRT; 377 AA.
AC Q9706; P78400;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE NATURAL KILLER CELL RECEPTOR.
GN KIR OR KIR-103AS
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97102173.
RA SELVAKUMAR A., STEFFENS U., DUPONT B.;
RT "NK cell receptor gene of the KIR family with two IG domains but
RT highest homology to KIR receptors with three IG domains.";
RL TISSUE ANTIGENS 48:285-294(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]

RA SELVAKUMAR A.;
RT "Polymorphism and domain variability of human killer cell inhibitory
RL IMMUNOL. REV. 155:183-196(1997).
DR EMBL; U7394; G1890666; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 360 AA; 39740 MW; B9C85111 CRC32;

Query Match      29.2%; Score 7; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.69e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYK 125
QY 10 VTGLYK 16
|||||

RESULT 5
ID O14622 PRELIMINARY; PRT; 342 AA.
AC O14622;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KILLER CELL RECEPTOR.
GN KIR103
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA SELVAKUMAR A., STEFFENS U., PALANISAMY N., CHAGANTI R.S.K.,
RA DUPONT B.;
RL TISSUE ANTIGENS 0:0-0(1997).
DR EMBL; AF002980; G2443480; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 342 AA; 37396 MW; 5C1B1963 CRC32;

Query Match      29.2%; Score 7; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.69e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYK 125
QY 10 VTGLYK 16
|||||

RESULT 6
ID Q99559 PRELIMINARY; PRT; 360 AA.
AC Q99559;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NK RECEPTOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOID;
RA BIASONI R.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X99479; E259274; -.
DR PFAM; PF00047; ig; 1.
SQ SEQUENCE 360 AA; 39702 MW; 1C5D2A07 CRC32;

Query Match      29.2%; Score 7; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.69e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYK 125
QY 10 VTGLYK 16
|||||

RESULT 7
ID P78401 PRELIMINARY; PRT; 360 AA.
AC P78401;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NK-RECEPTOR.
GN KIR-103AST.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97213129.
```

RP SEQUENCE FROM N.A.  
 RA SELVAKUMAR A., STEFFENS U., PALANISAMY N., CHAGANTI R.S.K.,  
 RA DUPONT B.;  
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; U71199, G1890662; -  
 DR EMBL; AF003123; G2228791; -  
 DR EMBL; AF003116; G2228791; JOINED.  
 DR EMBL; AF003117; G2228791; JOINED.  
 DR EMBL; AF003118; G2228791; JOINED.  
 DR EMBL; AF003119; G2228791; JOINED.  
 DR EMBL; AF003121; G2228791; JOINED.  
 DR EMBL; AF003122; G2228791; JOINED.  
 DR EMBL; AF003120; G2228791; JOINED.  
 DR PFAM; PF00047; 1g; 1.  
 SQ SEQUENCE 377 AA; 41527 MW; EF18859E CRC32;

Query Match 29.2%; Score 7; DB 4; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYEK 125  
 |||||  
 Qy 10 VTGLYEK 16

RESULT 10  
 ID Q99560 PRELIMINARY; PRT; 377 AA.  
 AC Q99560;  
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE NK RECEPTOR.  
 GN KIR2DL4.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
 OC CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LYMPHOID;  
 RA BIASSONI R.;  
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA VALIANT N.M., UHRBERG M., SHILLING H.G., LIENERT-WEIDENBACH K.,  
 RA ARNETT K.L., D'ANDREA A., PHILLIPS J.H., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA UHRBERG M., VALIANT N.M., SHUM B., SHILLING H.G.,  
 RA LIENERT-WEIDENBACH K., CORLISS B., TYAN D., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1998).  
 DR EMBL; X99480; E259275; -  
 DR EMBL; AF034771; G2739178; -  
 DR PFAM; PF00047; 1g; 1.  
 SQ SEQUENCE 377 AA; 41489 MW; D54A1B6A CRC32;

Query Match 29.2%; Score 7; DB 4; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYEK 125  
 |||||  
 Qy 10 VTGLYEK 16

RESULT 11  
 ID O43534 PRELIMINARY; PRT; 377 AA.  
 AC O43534;  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
 DE NATURAL KILLER CELL INHIBITORY RECEPTOR.  
 GN KIR2DL4.

OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
 OC CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA VALIANT N.M., UHRBERG M., SHILLING H.G., LIENERT-WEIDENBACH K.,  
 RA ARNETT K.L., D'ANDREA A., PHILLIPS J.H., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA UHRBERG M., VALIANT N.M., SHUM B., SHILLING H.G.,  
 RA LIENERT-WEIDENBACH K., CORLISS B., TYAN D., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1998).  
 DR EMBL; AF034773; G2739182; -  
 SQ SEQUENCE 377 AA; 41425 MW; 5C3C6A8D CRC32;

Query Match 29.2%; Score 7; DB 4; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYEK 125  
 |||||  
 Qy 10 VTGLYEK 16

RESULT 12  
 ID Q99562 PRELIMINARY; PRT; 377 AA.  
 AC Q99562;  
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE NK RECEPTOR.  
 GN KIR2DL4.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
 OC CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-LYMPHOID;  
 RA BIASSONI R.;  
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA VALIANT N.M., UHRBERG M., SHILLING H.G., LIENERT-WEIDENBACH K.,  
 RA ARNETT K.L., D'ANDREA A., PHILLIPS J.H., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA UHRBERG M., VALIANT N.M., SHUM B., SHILLING H.G.,  
 RA LIENERT-WEIDENBACH K., CORLISS B., TYAN D., LANIER L.L., PARHAM P.;  
 RL IMMUNITY 0:0-0(1998).  
 DR EMBL; X97229; E242794; -  
 DR EMBL; AF034772; G2739180; -  
 DR PFAM; PF00047; 1g; 1.  
 SQ SEQUENCE 377 AA; 41485 MW; 2B5EC7C5 CRC32;

Query Match 29.2%; Score 7; DB 4; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 119 VTGLYEK 125  
 |||||  
 Qy 10 VTGLYEK 16

RESULT 13  
 ID P78403 PRELIMINARY; PRT; 382 AA.  
 AC P78403;  
 DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
 DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE NK-RECEPTOR.  
 GN KIR-123FM.

OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: PRIMATES;  
OC CATARRHINI: HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 97213129.  
RA SELVAKUMAR A.;  
RT "Polymorphism and domain variability of human killer cell inhibitory  
receptors".  
RL IMMUNOL. REV. 155:183-196(1997).  
DR EMBL; U73396; G1890670; -.  
DR PFAM; PF00047; ig; 2.  
SQ SEQUENCE 382 AA; 42401 MW; 09E2368A CRC32;

Query Match 29.2%; Score 7; DB 4; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 VTGLYEK 223  
|||||  
QY 10 VTGLYEK 16

RESULT 14  
ID Q99564 PRELIMINARY; PRT; 382 AA.  
AC Q99564;  
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE NK RECEPTOR.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: PRIMATES;  
OC CATARRHINI: HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LYMPHOID;  
RA BIASSONI R.;  
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; X97233; E242797; -.  
DR PFAM; PF00047; ig; 2.  
SQ SEQUENCE 382 AA; 42446 MW; 046501BF CRC32;

Query Match 29.2%; Score 7; DB 4; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 VTGLYEK 223  
|||||  
QY 10 VTGLYEK 16

RESULT 15  
ID O43469 PRELIMINARY; PRT; 382 AA.  
AC O43469;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE NATURAL KILLER CELL RECEPTOR KIR3DS1 VARIANT.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA: PRIMATES;  
OC CATARRHINI: HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA VALIANT N.M., UHRBERG M., SHILLING H.G., LIENERT-WEIDENBACH K.,  
RA ARNETT K.L., D-ANDREA A., PHILLIPS J.H., LANIER L.L., PARHAM P.;  
RL IMMUNITY 0:0-0(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA UHRBERG M., VALIANT N.M., SHUM B., SHILLING H.G.,  
RA LIENERT-WEIDENBACH K., CORLISS B., TVAN D., LANIER L., PARHAM P.;  
RL IMMUNITY 0:0-0(1998).  
DR EMBL; AF022044; G2760895; -.  
SQ SEQUENCE 382 AA; 42446 MW; B6F84A2A CRC32;

Query Match 29.2%; Score 7; DB 4; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1.69e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 217 VTGLYEK 223  
|||||  
QY 10 VTGLYEK 16

Search completed: Sat Aug 28 14:43:39 1999  
Job time : 47 secs.

**This Page Blank (uspto)**

\*\*\*\*\*

WQSEFH (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:14:12 1999; MasPar time 4.68 Seconds

Tabular output not generated. 131.688 Million cell updates/sec

Title: >US-09-049-696-46  
Description: (1-29) from US09049696.pap  
Perfect Score: 186  
Sequence: 1 ITVTSKINKDTSKPSPLVYVNIROGAS 29

Scoring table: PAM 150  
Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 21.394; Variance 69.337; scale 0.309

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	64	34.4	572	36	Aspergillus fumigatus	3.23e+01
2	61	32.8	93	4	Non-A non-B hepatitis	6.33e+01
3	61	32.8	246	6	HCV NS4-NS5 peptide O	6.33e+01
4	60	32.3	175	2	Non-A non-B hepatitis	7.90e+01
5	60	32.3	184	2	Non-A non-B hepatitis	7.90e+01
6	60	32.3	246	6	HCV NS4-NS5 peptide O	7.90e+01
7	60	32.3	246	6	HCV NS4-NS5 peptide O	7.90e+01
8	60	32.3	246	6	HCV NS4-NS5 peptide O	7.90e+01
9	60	32.3	334	6	HCV NS4-NS5 peptide 1	7.90e+01
10	60	32.3	606	6	HCV NS4-NS5 peptide 1	7.90e+01
11	60	32.3	863	6	HCV NS4-NS5 peptide 1	7.90e+01
12	60	32.3	1411	6	HCV NS4-NS5 peptide 1	7.90e+01
13	60	32.3	2510	6	HCV antigen T7N1-30.	9.85e+01
14	59	31.7	21	24	50K-cellulase peptide	9.85e+01
15	59	31.7	3639	8	ACVS.	9.85e+01
16	59	31.7	3712	3	ACV synthetase.	9.85e+01

17	58	31.2	117	31	W28036	Staphylococcus aureus	1.23e+02
18	58	31.2	392	12	R62868	Korean-type Hepatitis	1.23e+02
19	58	31.2	433	25	W30256	Zuotin.	1.23e+02
20	58	31.2	449	29	W37129	Hepatitis C virus opt	1.23e+02
21	58	31.2	3010	6	R30616	Polyptide coded by	1.23e+02
22	58	31.2	3010	10	R53417	Blood transmissible N	1.23e+02
23	58	31.2	3011	7	R34468	Encoded by full-length	1.23e+02
24	57	30.6	452	24	W16544	50K-cellulase B from	1.52e+02
25	57	30.6	507	20	W01499	70 kD cellobiohydrola	1.52e+02
26	57	30.6	578	30	W40046	Biologically equivalent	1.52e+02
27	57	30.6	578	30	W40050	Biologically equivalent	1.52e+02
28	57	30.6	578	30	W40047	Biologically equivalent	1.52e+02
29	57	30.6	578	30	W40048	Biologically equivalent	1.52e+02
30	57	30.6	578	30	W40045	Biologically equivalent	1.52e+02
31	57	30.6	578	30	W40043	Biologically equivalent	1.52e+02
32	57	30.6	578	30	W40044	Biologically equivalent	1.52e+02
33	57	30.6	578	30	W40049	Biologically equivalent	1.52e+02
34	57	30.6	603	30	W40040	A Haemophilus influen	1.52e+02
35	57	30.6	650	4	R25270	Human B-raf protein k	1.52e+02
36	57	30.6	1070	4	R21521	Alpha galactosidase.	1.52e+02
37	56	30.1	204	5	R25890	HK15.	1.89e+02
38	56	30.1	204	4	R20986	Non-A, non-B hepatitis	1.89e+02
39	56	30.1	585	8	R41359	Tumour associated 90K	1.89e+02
40	56	30.1	585	8	R40172	Sequence of a 90K tum	1.89e+02
41	56	30.1	585	8	R40172	Sequence of an immuno	1.89e+02
42	56	30.1	585	7	R36533	GP85-97 clone 18 prod	1.89e+02
43	56	30.1	3010	18	R82694	Partial HCV non-struct	1.89e+02
44	56	30.1	3010	13	R68622	HCV protein cleavable	1.89e+02
45	56	30.1	3010	13	R68622		

ALIGNMENTS

RESULT 1  
ID W69391 standard; Protein; 572 AA.

AC W69391;  
DT 24-DEC-1998 (first entry)  
DE Aspergillus fumigatus protein 2  
KW ss: auxotrophic cell line; histidine; adenylic acid; leucine; growth;  
OS Aspergillus fumigatus.  
PN W09841621-A1.  
PD 24-SEP-1998.  
PF 18-MAR-1998; U05350.  
PR 18-MAR-1997; US-041300.  
PI (MILL-) MILLENNIUM PHARM INC.  
PI Gavrias V;  
DR WPI: 98-521216/44.  
DR N-PSDB: V58698.  
PT Aspergillus fumigatus polynucleotide(s) important for growth and  
PT reproduction - and auxotroph(s) created by omitting  
PT polynucleotide(s), useful e.g. to test functionality of unknown  
PT function cDNA from A. fumigatus cDNA library  
PS Claim 1; Fig 2: 53pp; English.  
CC The omission of one or more polynucleotides from Aspergillus fumigatus  
CC results in auxotrophic cell lines, which require media supplemented with  
CC histidine, adenylic acid or leucine respectively for growth and  
CC reproduction. The auxotrophic cell lines (especially which require  
CC histidine, adenylic acid or leucine to grow and reproduce) and  
CC polynucleotides can be used to test the functionality of unknown function  
CC cDNA from an A. fumigatus cDNA library. The polynucleotides and  
CC polypeptides are useful to identify agonists which may enhance growth  
CC and/or reproduction of A. fumigatus e.g. in the fermentation industry;  
CC they may also be administered (e.g. by inclusion of multiple gene copies)  
CC to enhance such growth and/or reproduction. The polypeptides can be used  
CC to produce antibodies, useful to detect polypeptides, screen for similar  
CC polypeptides from other organisms and as antimicrobials.  
SQ Sequence 572 AA;

Query Match 34.4%; Score 64; DB 36; Length 572;  
Best Local Similarity 38.5%; Pred. No. 3.23e+01;  
Matches 10; Conservative 7; Mismatches 8; Indels 1; Gaps 1;



CC human liver tissue or serum is used to synthesize cDNA. A library  
 CC is constructed using a lambda gtl1 cloning system. This library  
 CC is screened and cloned for antigenic activity. Clones isolated  
 CC can be incorporated into a vector plasmid, which is then inserted  
 CC into E.coli to give a transformant which expressed an antigenic  
 CC protein having this amino acid sequence.  
 CC The product may be used for the immunoassay of antibodies to NANBH  
 CC antigen in samples of serum etc. for diagnostic purposes.  
 CC See also Q10523-540.  
 CC Sequence 175 AA;

Query Match 32.3%; Score 60; DB 2; Length 175;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 149 isvaaeilkrkfpalpiwa 170

QY 1 ITVTSKTKDTSKFPSPLVVYA 22

## RESULT 5

ID R10750 standard; Protein; 184 AA.

AC R10750;

DT 22-APR-1991 (first entry)

DE Non-A non-B hepatitis specific antigenic protein encoded by phage  
 clone lambda HC2248.

KW Non-A non-B hepatitis; antigenic protein; NANBH; phage clone;

KW Immunoassay; antibodies; diagnosis.

OS Homo sapiens.

PN WO9101376-A.

PD 07-FEB-1991.

PF 13-JUL-1990; J00906.

PR 14-JUL-1989; JP-182073.

PR 19-JUL-1989; JP-184739.

PR 22-JUL-1989; JP-189874.

PR 27-JUL-1989; JP-192721.

PR 29-JUL-1989; JP-195413.

PR 03-AUG-1989; JP-200217.

PR 10-AUG-1989; JP-205722.

PR 21-SEP-1989; JP-243304.

PR 22-SEP-1989; JP-245268.

PR 19-OCT-1989; JP-270398.

PA (CHUS ) CHUGAI SEIYAKU KK.

PI Arima T, Yamamoto O, Tsuchiya M, Oshima M;

DR WPI: 91-058149/08.

DR N-PSDB; Q10531.

PT Antigenic protein specific for non-A, non-B hepatitis - and cDNA  
 coding for it which corresponds to RNA of infected liver tissue

PT or serum

PS Disclosure; Fig 11; 69pp; Japanese.

CC The DNA corresponding to an RNA isolated directly from infected  
 CC human liver tissue or serum is used to synthesize cDNA. A library  
 CC is constructed using a lambda gtl1 cloning system. This library  
 CC is screened and cloned for antigenic activity. Clones isolated  
 CC can be incorporated into a vector plasmid, which is then inserted  
 CC into E.coli to give a transformant which expressed an antigenic  
 CC protein having this amino acid sequence.

CC The product may be used for the immunoassay of antibodies to NANBH  
 CC antigen in samples of serum etc. for diagnostic purposes.

CC See also Q10523-540.

CC Sequence 184 AA;

Query Match 32.3%; Score 60; DB 2; Length 184;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 149 isvaaeilkrkfpalpiwa 170

QY 1 ITVTSKTKDTSKFPSPLVVYA 22

## RESULT 6

ID R29894 standard; Protein; 246 AA.

R29894;  
 AC 26-APR-1993 (first entry)  
 DE HCV NS4-NS5 peptide Q28-4.  
 KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;  
 CC transcriptase; cDNA; primer; allele; core; region; upstream;  
 CC hydrophilic; turn structure; alpha helix; beta sheet; antigen;  
 KW determinant; antiserum.  
 OS Hepatitis C virus  
 PN EP-518313-A.  
 PD 16-DEC-1992.  
 PF 11-JUN-1992; 109812.  
 PR 11-JUN-1991; JP-139268.  
 PR 12-JUL-1991; JP-172794.  
 PR 07-OCT-1991; JP-287008.  
 PR 16-DEC-1991; JP-332329.  
 PR 20-APR-1992; JP-099957.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;  
 PI Teranishi Y;  
 DR WPI: 92-417213/51.  
 DR N-PSDB; Q32525.  
 PT New hepatitis C virus gene and its encoded protein - used for  
 PT diagnosing and vaccinating against hepatitis C virus infections  
 PS Disclosure; Page 240-42; 305pp; English.  
 CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.  
 CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 246 AA;

Query Match 32.3%; Score 60; DB 6; Length 246;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 19 vsvaaeilkrkfpampywa 40

QY 1 ITVTSKTKDTSKFPSPLVVYA 22

## RESULT 7

ID R29892 standard; Protein; 246 AA.

AC R29892;

DT 26-APR-1993 (first entry)

DE HCV NS4-NS5 peptide Q28-1.

KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele; core; region; upstream;

KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;

OS Hepatitis C virus

PN EP-518313-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 109812.

PR 11-JUN-1991; JP-139268.

PR 12-JUL-1991; JP-172794.

PR 07-OCT-1991; JP-287008.

PR 16-DEC-1991; JP-332329.

PR 20-APR-1992; JP-099957.

PA (MITU ) MITSUBISHI KASEI CORP.

PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;

PI Teranishi Y;

DR WPI: 92-417213/51.

DR N-PSDB; Q32523.

PT New hepatitis C virus gene and its encoded protein - used for

PT diagnosing and vaccinating against hepatitis C virus infections  
 PS Disclosure: Page 237-38; 305pp; English.  
 CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.  
 CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 246 AA;

Query Match 32.3%; Score 60; DB 6; Length 246;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 19 vsvaaeilrktkfkpaampvwa 40  
 QY 1 IIVTSKTNKDTSKFSPPLVYVA 22

## RESULT 8

ID R29893 standard; Protein; 246 AA.

AC R29893; (first entry)  
 DE HCV NS4-NS5 peptide Q28-2.  
 KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;  
 KW transcriptase; cDNA; primer; allele; core; region; upstream;  
 KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;  
 KW determinant; antiserum.

OS Hepatitis C virus

PN EP-518313-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 109812.

PR 11-JUN-1991; JP-139268.

PR 12-JUL-1991; JP-172794.

PR 07-OCT-1991; JP-287008.

PR 16-DEC-1991; JP-332329.

PR 20-APR-1992; JP-099957.

PA (MITU) MITSUBISHI KASEI CORP.

PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;

PI Teranishi Y;

DR WPI; 92-417213/51.

DR N-PSDB; Q32524.

PT New hepatitis C virus gene and its encoded protein - used for

PT diagnosing and vaccinating against hepatitis C virus infections

PS Disclosure: Page 239-40; 305pp; English.

CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.

CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 246 AA;

Query Match 32.3%; Score 60; DB 6; Length 246;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 19 vsvaaeilrktkfkpaampvwa 40  
 QY 1 IIVTSKTNKDTSKFSPPLVYVA 22

## RESULT 9

ID R29878 standard; Protein; 334 AA.

AC R29878;

DE HCV NS4-NS5 peptide 1728.

KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele; core; region; upstream;

KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;

KW determinant; antiserum.

OS Hepatitis C virus

PN EP-518313-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 109812.

PR 11-JUN-1991; JP-139268.

PR 12-JUL-1991; JP-172794.

PR 07-OCT-1991; JP-287008.

PR 16-DEC-1991; JP-332329.

PR 20-APR-1992; JP-099957.

PA (MITU) MITSUBISHI KASEI CORP.

PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;

PI Teranishi Y;

DR WPI; 92-417213/51.

DR N-PSDB; Q32509.

PT New hepatitis C virus gene and its encoded protein - used for

PT diagnosing and vaccinating against hepatitis C virus infections

PS Disclosure: Page 204-06; 305pp; English.

CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.

CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 334 AA;

Query Match 32.3%; Score 60; DB 6; Length 334;

Best Local Similarity 31.8%; Pred. No. 7.90e+01;

Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 107 vsvaaeilrktkfkpaampvwa 128  
 QY 1 IIVTSKTNKDTSKFSPPLVYVA 22

## RESULT 10

ID R29880 standard; Protein; 606 AA.

AC R29880;

DE HCV NS4-NS5 peptide 1718.

KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele; core; region; upstream;

KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;

KW determinant; antiserum.

OS Hepatitis C virus

PN EP-518313-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 109812.

PR 11-JUN-1991; JP-139268.

FT misc\_difference 326

FT /label= Leu



PR 12-JUL-1991; JP-172794.  
 PR 07-OCT-1991; JP-287008.  
 PR 16-DEC-1991; JP-332329.  
 PR 20-APR-1992; JP-099957.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;  
 PI Teranishi Y;  
 DR WPI; 92-417213/51.  
 DR N-PSDB; Q32511.  
 PT New hepatitis C virus gene and its encoded protein - used for  
 PT diagnosing and vaccinating against hepatitis C virus infections  
 PS Disclosure; Page 208-11; 305pp; English.  
 CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.  
 CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 606 AA;

Query Match 32.3%; Score 60; DB 6; Length 606;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 107 vsaaeilrkrtrkfpampywa 128  
 :|:: :|::|::|  
 QY 1 ITVTSKTKNTSRFPSPPLVYA 22

RESULT 11  
 ID R29881 standard; Protein; 863 AA.

AC R29881;  
 DT 26-APR-1993 (first entry)  
 DE HCV NS4-NS5 peptide 2218.  
 KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;  
 KW transcriptase; cDNA; primer; allele; core; region; upstream;  
 KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;  
 KW determinant; antiserum.  
 OS Hepatitis C virus  
 FH Key Location/Qualifiers  
 FT misc\_difference 583  
 FT /label= Leu

PN EP-518313-A.  
 PD 16-DEC-1992.  
 PF 11-JUN-1992; 109812.  
 PR 11-JUN-1991; JP-139268.  
 PR 12-JUL-1991; JP-172794.  
 PR 07-OCT-1991; JP-287008.  
 PR 16-DEC-1991; JP-332329.  
 PR 20-APR-1992; JP-099957.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;  
 PI Teranishi Y;  
 DR WPI; 92-417213/51.  
 DR N-PSDB; Q32512.  
 PT New hepatitis C virus gene and its encoded protein - used for  
 PT diagnosing and vaccinating against hepatitis C virus infections  
 PS Disclosure; Page 212-16; 305pp; English.  
 CC The sequences given in R29871-906 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated

CC represent different alleles of the same region of the HCV gene.  
 CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 863 AA;

Query Match 32.3%; Score 60; DB 6; Length 863;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 364 vsaaeilrkrtrkfpampywa 385  
 :|:: :|::|::|  
 QY 1 ITVTSKTKNTSRFPSPPLVYA 22

RESULT 12  
 ID R29533 standard; Protein; 1411 AA.

AC R29533;  
 DT 26-APR-1993 (first entry)  
 DE HCV NS4-NS5 peptide 1530U.  
 KW Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;  
 KW transcriptase; cDNA; primer; allele; core; region; upstream;  
 KW hydrophilic; turn structure; alpha helix; beta sheet; antigen;  
 KW determinant; antiserum.  
 OS Hepatitis C virus  
 PN EP-518313-A.  
 PD 16-DEC-1992.  
 PF 11-JUN-1992; 109812.  
 PR 11-JUN-1991; JP-139268.  
 PR 12-JUL-1991; JP-172794.  
 PR 07-OCT-1991; JP-287008.  
 PR 16-DEC-1991; JP-332329.  
 PR 20-APR-1992; JP-099957.  
 PA (MITU ) MITSUBISHI KASEI CORP.  
 PI Hayashi N, Honda Y, Murakami T, Seki M, Takahashi K;  
 PI Teranishi Y;  
 DR WPI; 92-417213/51.  
 DR N-PSDB; Q32443.

PT New hepatitis C virus gene and its encoded protein - used for  
 PT diagnosing and vaccinating against hepatitis C virus infections  
 PS Disclosure; Page 216-24; 305pp; English.  
 CC The sequences given in R29871-81 and R29533 are encoded by various  
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of  
 CC the invention. These NS4-NS5 RNA sequences were isolated from the  
 CC serum of a patient suffering from hepatitis C (HC). The isolated RNA  
 CC sequences were converted into cDNA using transcriptase in the presence  
 CC of one of the primer sequences given in Q32565-77. The sequences were  
 CC then amplified using primer pairs. The cDNA sequences isolated  
 CC represent different alleles of the same region of the HCV gene.  
 CC Sequence analysis shows that these clones represent the core region  
 CC and some upstream sequences of HCV. These polypeptides are thought to  
 CC contain a highly hydrophilic region which can adopt a "turn structure"  
 CC which is not an alpha helix or a beta sheet. These polypeptides are  
 CC thought to act as antigen determinants and are highly reactive with  
 CC antiserum raised against HCV-associated antigens. See also Q32436.  
 SQ Sequence 1411 AA;

Query Match 32.3%; Score 60; DB 6; Length 1411;  
 Best Local Similarity 31.8%; Pred. No. 7.90e+01;  
 Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Db 669 vsaaeilrkrtrkfpampywa 690  
 :|:: :|::|::|  
 QY 1 ITVTSKTKNTSRFPSPPLVYA 22

RESULT 13  
 ID R29527 standard; Protein; 2510 AA.

AC R29527; 1993 (first entry)  
 DT 26-APR-1993

```

PT useful for bio-stoning and bio-finishng cotton fabrics, especially
PT denim, and as detergent for treating cellulosic fibre containing
PT textiles
PS Example 11; Page 53; 204pp; English.
PS 12 Peptides (W16546-57) were generated from an isolated 50 kDa
PS cellulase of Melanocarpus albomyces ALK04237 by treatment of the
PS enzyme with tyranogen bromide, and digestion of the resulting
PS fragments with trypsin or lysylendopeptidase C. Peptides 507
PS (W16547) and 509 (W16547) were used to design primers (see
PS W65547-48) utilised in the PCR amplification of a gene (T66542)
PS coding for the 50K-cellulase (W16543). The enzyme can be used
PS for bio-stoning and bio-finishng cotton fabrics.
PS Sequence 21 AA;
SQ
Query Watch 31.7%; Score 59; DB 24; Length 21;
Best Local Similarity 61.5%; Pred. No. 9.85e+01;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps
0;
Db 3 pspevtysnlrxg 15
||| | | | | |
QY 15 PSEPLVYANIRQG 27
||| | | | | |
RESULT 15
ID R40227 standard; Protein: 3639 AA.
AC R40227;
DE 21-FEB-1994 (first entry)
DT ACVS.
KW Delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase; ACVS;
OS beta-lactam; antibiotic; transformed; cephalosporin; vector.
SS Acronemium chrysogenum.
PN J05192162-A.
PD 03-AUG-1993.
PF 25-JUL-1991; 186222.
PR 31-JUL-1990; JP-205677.
PA (TAKE ) TAKEDA CHEM IND LTD.
PE WPI; 93-2777475/35.
DR N-PSDB; Q48231.
DR DNA coding delta-(L-alpha-amino-adipyl) L-cystinyl D-valine
PT synthase - for improved productivity of cephalosporin antibiotics
PS Claim 1; Page 14-27; 69pp; Japanese.
PS The sequence (Q48231) is of a vector which includes the
PS delta-(L-alpha-aminoadipyl)-L-cystinyl-D-valine synthase gene.
PS This sequence was transformed into a host cell to express the ACVS
PS product. The protein produced (R40227) was then used to manufacture
PS a beta-lactam antibiotic.
PS Sequence 3639 AA;
SQ
Query Watch 31.7%; Score 59; DB 8; Length 3639;
Best Local Similarity 27.8%; Pred. No. 9.85e+01;
Matches 5; Conservative 7; Mismatches 6; Indels 0; Gaps
0;
Db 1029 nearyskvpladydvre 1046
:::|:|:| | | | |
QY 9 KDTSKFPSPLVYANIRQ 26
:::|:|:| | | | |
Search completed: Mon Aug 23 13:14:43 1999
Job time : 31 secs.

```

\*\*\*\*\*

W A R E H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:13:43 1999; MasPar time 6.15 Seconds

Tabular output not generated.  
188.945 Million cell updates/sec

Title: >US-049-696-46  
Description: (1-29) from US09049696.pep  
Perfect Score: 186  
Sequence: 1 ITVTSKNTKTSKFPSPLVVYNIROGAS 29

Scoring table: PAM 150  
Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 29.792; Variance 42.775; scale 0.696

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	143	76.9	913	3	gob-5 protein - Mouse	2.17e-17
2	94	50.5	794	2	hypothetical protein	4.13e-06
3	94	50.5	820	2	Lu-ECAM-1 protein - b	4.13e-06
4	94	50.5	905	2	Lu-ECAM-1 protein - b	4.13e-06
5	75	40.3	537	2	cellulose 1,4-beta-ce	2.97e-02
6	72	38.7	394	2	hypothetical protein	1.10e-01
7	71	38.2	147	2	conserved hypothetical	1.69e-01
8	69	37.1	976	2	pleiotropic drug resi	3.93e-01
9	67	36.0	89	1	phosphotransferase sy	9.02e-01
10	65	34.9	229	2	F37A4_3 protein - Cae	2.04e-00
11	65	34.9	506	2	hypothetical protein	2.04e-00
12	64	34.4	960	2	discs-large tumor sup	3.04e-00
13	63	33.9	523	2	hypothetical 56.4 kD	4.52e-00
14	63	33.9	4436	2	hypothetical protein	4.52e-00
15	62	33.3	449	2	GTP-binding regulator	6.70e-00
16	62	33.3	449	2	probable phosphomanno	6.70e-00
17	62	33.3	780	2	hypothetical protein	6.70e-00
18	61	32.8	329	2	RNA-binding protein c	9.86e-00
19	61	32.8	1745	2	tight junction-associ	9.86e-00
20	60	32.3	167	2	hypothetical protein	1.45e-01
21	60	32.3	175	2	polypeptide (clone 22	1.45e-01
22	60	32.3	182	2	polypeptide (clone 22	1.45e-01
23	60	32.3	183	2	hypothetical 20.5 kD	1.45e-01

24	60	32.3	333	2	H71703	threonine dehydratase	1.45e-01
25	60	32.3	665	2	S28719	gene g-1 protein - sl	1.45e-01
26	59	31.7	178	2	S03216	hypothetical protein	2.11e-01
27	59	31.7	586	1	HMVNI1	hemagglutinin precurs	2.11e-01
28	59	31.7	567	2	S49600	glucose transport pro	2.11e-01
29	59	31.7	764	2	H71607	hypothetical protein	2.11e-01
30	59	31.7	1122	2	S64443	probable membrane pro	2.11e-01
31	59	31.7	3712	1	YGCEVC	alpha-aminoadipyl-cys	2.11e-01
32	58	31.2	116	2	S57106	probable membrane pro	3.06e-01
33	58	31.2	261	2	T02107	hypothetical protein	3.06e-01
34	58	31.2	277	2	B69479	conserved hypothetical	3.06e-01
35	58	31.2	426	2	E70368	histidinol dehydrogen	3.06e-01
36	58	31.2	433	2	S25194	zuotin - yeast (Sacch	3.06e-01
37	58	31.2	570	1	RDECSH	sulfite reductase (NA	3.06e-01
38	58	31.2	944	2	S48921	probable membrane pro	3.06e-01
39	58	31.2	1016	2	T00375	KIAA0647 protein - hu	3.06e-01
40	58	31.2	1024	2	S71804	receptor-like serpent	3.06e-01
41	58	31.2	1778	2	JT0382	apolipoprotein B - pi	3.06e-01
42	58	31.2	2629	2	I46569	apolipoprotein B - pi	3.06e-01
43	58	31.2	3010	1	GNWVTV	genome polyprotein -	3.06e-01
44	58	31.2	3010	2	SI8030	genome polyprotein -	3.06e-01
45	58	31.2	3010	2	A45573	polyprotein(protein N	3.06e-01

ALIGNMENTS

RESULT 1 JG0168 #type complete  
ENTRY gob-5 protein - Mouse  
TITLE #formal\_name Mus musculus #common\_name house mouse  
ORGANISM 18-Mar-1999 #sequence\_revision 18-Mar-1999 #text\_change  
DATE 18-Mar-1999  
ACCESSIONS JG0168  
REFERENCE #authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.  
#journal Biochem. Biophys. Res. Commun. (1995) 255:347-351  
#title Cloning and identification of the gene gob-5, which is  
#accession expressed in intestinal goblet cells in mice.  
#status preliminary  
#residues 1-913 ##label KOM  
##cross-references DBJ:AB016592  
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755  
Query Match 76.9%; Score 143; DB 3; Length 913;  
Best Local Similarity 72.4%; Pred. No. 2.17e-17;  
Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 592 ITVTPVVKNTGKFPSPVTVYASIROGAS 620  
QY 1 ITVTSKNTKTSKFPSPLVVYNIROGAS 29  
||||| :|:|:||||| :|:|:|||||

RESULT 2 T02171 #type complete  
ENTRY hypothetical protein Lu-ECAM-1 - bovine  
TITLE #formal\_name Bos primigenius taurus #common\_name cattle  
ORGANISM 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change  
DATE 05-Mar-1999  
ACCESSIONS T02171  
REFERENCE Z14590  
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;  
Levine, R.; Goodwin, A.; Pauli, B.O.  
#submission submitted to the EMBL Data Library, April 1997  
#description Cloning and characterization of Lu-ECAM-1 suggest it is an  
endothelial chloride channel.  
#accession T02171  
#status preliminary; translated from GB/EMBL/DBJ  
##molecule\_type mRNA  
##residues 1-794 ##label ELB  
##cross-references EMBL:AF001262; NID:g2623764; PID:g2623765  
##experimental\_source lung  
SUMMARY #length 794 #molecular-weight 89509 #checksum 6853

```
Query Match          50.5%; Score 94; DB 2; Length 794;
Best Local Similarity 37.0%; Pred. No. 4.13e-06;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAHMSQHTAHYPSPMIVYAOVSQ 625
QY 1 ITVTSKTKNDTSKFPSPLVVYANIROG 27

RESULT 3
ENTRY T02152 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02152
REFERENCE Z14590
#authors Eible, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.U.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
#description endothelial chloride channel.
#accession T02152
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
##residues 1-820 ##label ELB
##cross-references EMBL:AF001263; NID:g2623766; PID:g2623767
##experimental_source lung
SUMMARY #length 820 #molecular-weight 91464 #checksum 3104

Query Match          50.5%; Score 94; DB 2; Length 820;
Best Local Similarity 37.0%; Pred. No. 4.13e-06;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAHMSQHTAHYPSPMIVYAOVSQ 625
QY 1 ITVTSKTKNDTSKFPSPLVVYANIROG 27

RESULT 4
ENTRY T02205 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02205
REFERENCE Z14590
#authors Eible, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.U.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
#description endothelial chloride channel.
#accession T02205
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
##residues 1-905 ##label ELB
##cross-references EMBL:AF001261; NID:g2623762; PID:g2623763
##experimental_source lung
SUMMARY #length 905 #molecular-weight 101005 #checksum 5557

Query Match          50.5%; Score 94; DB 2; Length 905;
Best Local Similarity 37.0%; Pred. No. 4.13e-06;
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAHMSQHTAHYPSPMIVYAOVSQ 625
QY 1 ITVTSKTKNDTSKFPSPLVVYANIROG 27

RESULT 5
ENTRY JU0150 #type complete
TITLE cellulose 1,4-beta-cellobiosidase (EC 3.2.1.91) I -
```

```
Penicillium janthinellum
exo-cellobiohydrolase
#formal_name Penicillium janthinellum
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
04-Sep-1998
ACCESSIONS JU0150; S21508
REFERENCE JU0150
#authors Christoph, A.K.; Weigel, T.O.; Schulz, G.
#journal Gene (1993) 124:57-65
#title Cloning, sequencing, and heterologous expression of a
#title cellulase-encoding cDNA (cbhl) from Penicillium
#accession JU0150
#molecule_type mRNA
##residues 1-537 ##label CHR
##cross-references EMBL:X59054; NID:g3177; PID:e49600; PID:g3178
##note the authors translated the codon ATG for residue 172 as
Asn

GENETICS cbhl
#gene #superfamily cellulose 1,4-beta-cellobiosidase I; fungal
CLASSIFICATION cellulose-binding domain homology
KEYWORDS glycosidase; hydrolase; polysaccharide degradation
FEATURE #domain fungal cellulose-binding domain homology #label
506-537 FCB
SUMMARY #length 537 #molecular-weight 56842 #checksum 9925

Query Match          40.3%; Score 75; DB 2; Length 537;
Best Local Similarity 30.8%; Pred. No. 2.97e-02;
Matches 8; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Db 429 ISRRENTVESTYPNAYVIYSNIKTGP 454
QY 3 VTSTKTKNDTSKFPSPLVVYANIROGA 28

RESULT 6
ENTRY E64028 #type complete
TITLE hypothetical protein HI1410 - Haemophilus influenzae (strain
Rd KW20)
ORGANISM #formal_name Haemophilus influenzae
DATE 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
10-Oct-1997
ACCESSIONS E64028
REFERENCE A64000
#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
Kirkness, E.F.; Kesteven, A.R.; Bult, C.J.; Tomb, J.F.;
Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
J.L.; Geoghegan, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus
#title influenzae Rd.
#cross-references MUID:95350630
#accession E64028
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
##residues 1-394 ##label TIGR
##cross-references GB:U32820; GB:L42023; NID:g1574231; PID:g1574247;
TIGR:HI1410

GENETICS GTG
#start_codon
SUMMARY #length 394 #molecular-weight 44782 #checksum 91

Query Match          38.7%; Score 72; DB 2; Length 394;
Best Local Similarity 31.0%; Pred. No. 1.10e-01;
Matches 9; Conservative 10; Mismatches 10; Indels 0; Gaps 0;
```

```

Db      226 IDSANRTNQSAVKFADLIIFDSIGVCG 254
      1 ITVTSKTNKDTSKPPSPPLVYANIROGAS 29

RESULT 7
ENTRY   #type complete
TITLE   conserved hypothetical protein Ar1962 - Archaeoglobus
        fulgidus
ORGANISM #formal_name Archaeoglobus fulgidus
DATE      05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
        05-Jun-1998
ACCESSIONS A69495
REFERENCE  A69250
#authors   Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
           K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.;
           Peterson, J.D.; Richardson, D.L.; Kervatage, A.R.; Graham,
           D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
           Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
           Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.;
           Peterson, S.; Reich, C.I.; McNeill, L.K.; Badger, J.H.;
           Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
           J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
           T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.;
           D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.;
           Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
           C.R.; Venter, J.C.
#journal   Nature (1997) 390:364-370
#title     The complete genome sequence of the hyperthermophilic,
           sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references MUID:98049343
#accession  A69495
#status    preliminary; nucleic acid sequence not shown;
           translation not shown
#molecule_type DNA
#residues  1-147 #label KLE
#cross-references GB:AE000967; GB:AE000782; NID:g2689290; PID:g2648577;
           TIGR:AF1962
SUMMARY   #length 147 #molecular-weight 16481 #checksum 6434

Query Match      38.28; Score 71; DB 2; Length 147;
Best Local Similarity 47.18; Pred. No. 1.69e-01;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db      75 SRFRPLIVYSADRHNA 91
      12 SKFPSPLVYANIROGA 28

RESULT 8
ENTRY   #type complete
TITLE   pleiotropic drug resistance protein PDR3 - yeast
        (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YBL005w; protein YBL0323
ORGANISM #formal_name Saccharomyces cerevisiae
DATE      26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change
        17-Mar-1999
ACCESSIONS S45738; S47186; S25328; S46661
REFERENCE  S45736
#authors   Delaveau, T.; Jacq, C.; Perea, J.
#submission submitted to the Protein Sequence Database, August 1994
#accession  S45738
#molecule_type DNA
#residues  1-976 #label DEL
#cross-references EMBL:Z35766; NID:g535983; PID:g535984; MIPS:YBL005w
#experimental_source strain S288C
REFERENCE  S47186
#authors   Delaveau, T.; Delahodde, A.; Carvajal, E.; Subik, J.; Jacq,
           C.
#submission submitted to the EMBL Data Library, June 1994
#description PDR3, a new yeast regulatory gene, is homologous to PDR1 and
           controls the multidrug resistance phenomenon.

```

```

#accession  S47186
#molecule_type DNA
#residues  1-976 #label DE2
#cross-references EMBL:X79803; NID:g506477; PID:g506478
#experimental_source strain S288C
REFERENCE  S25326
#authors   Delaveau, T.; Jacq, C.; Perea, J.
#journal   Yeast (1992) 8:761-768
#title     Sequence of a 12.7 kb segment of yeast chromosome II
           identifies a PDR-like gene and several new open reading
           frames.
#cross-references MUID:93070613
#accession  S25328
#molecule_type DNA
#residues  1-191 #label DE3
#cross-references EMBL:S47695
#experimental_source strain S288C
#note      the authors did not translate the codons for residues 1
           and 185-191
REFERENCE  S46661
#authors   Delaveau, T.; Delahodde, A.; Carvajal, E.; Subik, J.; Jacq,
           C.
#journal   Mol. Gen. Genet. (1994) 244:501-511
#title     PDR3, a new yeast regulatory gene, is homologous to PDR1 and
           controls the multidrug resistance phenomenon.
#cross-references MUID:94359476
#accession  S46661
#molecule_type DNA
#residues  1-976 #label DEF
#cross-references EMBL:X79803; NID:g506477; PID:g506478
GENETICS
#gene      SGD:PDR3
#cross-references SGD:S0000101; MIPS:YBL005w
#map_position 2L
CLASSIFICATION #superfamily unassigned GAL4-type zinc cluster proteins; GAL4
               zinc binuclear cluster homology
KEYWORDS      DNA binding; nucleus; transcription regulation; transmembrane
               protein; zinc finger
FEATURE
10-46         #domain GAL4 zinc binuclear cluster homology #label
               GAL4\
15-41         #region zinc finger\
701-720       #domain transmembrane #status predicted #label TM1\
773-789       #domain transmembrane #status predicted #label TM2
SUMMARY       #length 976 #molecular-weight 112570 #checksum 3454

Query Match      37.13; Score 69; DB 2; Length 976;
Best Local Similarity 56.33; Pred. No. 3.93e-01;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db      127 KDDSKYONQLVYONI 142
      9 KDTSKFPSPLVYANI 24

RESULT 9
ENTRY   #type complete
TITLE   phosphotransferase system phosphohistidine-containing protein
        - Mycoplasma capricolum (SGC3)
ALTERNATE_NAMES phosphocarrier protein Hpr
ORGANISM #formal_name Mycoplasma capricolum
DATE      26-May-1995 #sequence_revision 31-Jan-1997 #text_change
        09-Apr-1998
ACCESSIONS A49683
REFERENCE  A49683
#authors   Zhu, P.P.; Reizer, J.; Reizer, A.; Peterkofsky, A.
#journal   J. Biol. Chem. (1993) 268:26531-26540
#title     Unique monocistronic operon (ptsH) in Mycoplasma capricolum
           encoding the phosphocarrier protein, Hpr, of the
           phosphoenolpyruvate:sugar phosphotransferase system.
           Cloning, sequencing, and characterization of ptsH.
#cross-references MUID:94075343
#accession  A49683

```

```
#status      preliminary
#molecule_type DNA
#residues    1-89 #label ZHU
#cross-references GB:L22432; NID:g435095; PID:g435097
#note        monocistronic operon

GENETICS
#genetic_code SGC3
CLASSIFICATION #superfamily phosphotransferase system
                phosphohistidine-containing protein; phosphotransferase
                system phosphohistidine-containing protein homology
KEYWORDS       phosphocarrier protein; phosphohistidine; phosphoprotein;
                phosphotransferase system; sugar transport system

FEATURE
8-85           #domain phosphotransferase system
                phosphohistidine-containing protein homology #label
                HPR\
15             #binding_site phosphate (His) (covalent) (by
                phosphotransferase system enzyme I) #status predicted\
46             #binding_site phosphate (Ser) (covalent) #status
                predicted
SUMMARY        #length 89 #molecular-weight 9418 #checksum 3831
                Query Match      36.0%; Score 67; DB 1; Length 89;
                Best Local Similarity 47.4%; Pred. No. 9.02e-01;
                Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 24 KEAKFSNNITIIANEKQG 42
   |||||: : : || :||
QY 9 KOTSKFSPLVVYVNIROG 27

RESULT 10
ENTRY   S44640 #type complete
TITLE   F37A4.3 protein - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE     20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
        31-Oct-1997
ACCESSIONS S44640
REFERENCE   S44638
#authors    Fulton, L.
#submission submitted to the EMBL Data Library, February 1994
#description The sequence of C. elegans cosmid F37A4.
#accession  S44640
#status     preliminary
#molecule_type DNA
#residues   1-229 #label FUL
#cross-references EMBL:U00032; NID:g458958; PID:g458961

GENETICS
#introns    34/2; 55/3; 93/2
SUMMARY     #length 229 #molecular-weight 27604 #checksum 7683
                Query Match      34.9%; Score 65; DB 2; Length 229;
                Best Local Similarity 44.4%; Pred. No. 2.04e-00;
                Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 125 KLSKFDHSLPIYTIIRD 142
   ||| | | : : | :
QY 9 KOTSKFSPLVVYVNIROG 26

RESULT 11
ENTRY   S34497 #type complete
TITLE   hypothetical protein 506 (psbd 3' region) - Euglena gracilis
                Chloroplast
ORGANISM #formal_name chloroplast Euglena gracilis
DATE     30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
        09-Sep-1997
ACCESSIONS S34497; S34865
REFERENCE   S34494
#authors    Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort,
                A.; Orsat, B.; Spielmann, A.; Stutz, E.
#submission submitted to the EMBL Data Library, January 1993
#description The complete sequence of the Euglena gracilis chloroplast
```

```
#accession    genome (tentative).
#molecule_type DNA
#residues     1-506 #label HAL1
#cross-references EMBL:X70810; NID:g415327; PID:g295955
REFERENCE     S34862
#authors      Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.;
                Monfort, A.; Orsat, B.; Spielmann, A.; Stutz, E.
#journal      Nucleic Acids Res. (1993) 21:3537-3544
#title        Complete sequence of Euglena gracilis chloroplast DNA.
#cross-references MUID:93347989
#accession    S34865
#status       nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues     1-506 #label HAL2
#cross-references EMBL:X70810; NID:g415327; PID:g295955
#note         the nucleotide sequence was submitted to the EMBL Data
                Library, January 1993
GENETICS
#genome       chloroplast
KEYWORDS      chloroplast
SUMMARY       #length 506 #molecular-weight 59802 #checksum 8794
                Query Match      34.9%; Score 65; DB 2; Length 506;
                Best Local Similarity 53.8%; Pred. No. 2.04e-00;
                Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 366 KFSKPLVLFDNK 378
   ||: |||: ||:
QY 13 KFPSPLVYVNIIR 25

RESULT 12
ENTRY   A39651 #type complete
TITLE   discs-large tumor suppressor - fruit fly (Drosophila
                melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE     30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change
        24-Sep-1998
ACCESSIONS A39651
REFERENCE   A39651
#authors    Woods, D.F.; Bryant, P.J.
#journal     Cell (1991) 66:451-464
#title       The discs-large tumor suppressor gene of Drosophila encodes a
                guanylate kinase homolog localized at septate junctions.
#cross-references MUID:91330294
#accession  A39651
#molecule_type mRNA
#residues   1-960 #label WOO
#cross-references GB:M73529; NID:gl57243; PID:gl57244
COMMENT     Loss of this protein causes large imaginal disks by allowing
                neoplastic overgrowth. The presence of a guanylate kinase domain
                suggests that function of this protein involves signal
                transduction to control cellular proliferation.

GENETICS
#gene        FlyBase:dlgl1
#cross-references FlyBase:FBgn0001624
CLASSIFICATION #superfamily discs-large tumor suppressor; GLGF domain
                homology; guanylate kinase homology; SH3 homology
                signal transduction
KEYWORDS      signal transduction
FEATURE       45-123
                159-241
                491-563
                607-665
                771-948
SUMMARY       #length 960 #molecular-weight 102468 #checksum 8684
                Query Match      34.4%; Score 64; DB 2; Length 960;
                Best Local Similarity 33.3%; Pred. No. 3.04e-00;
                Matches 8; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

Db 346 VTATATASNDSSKLPPLGANSI 369
```

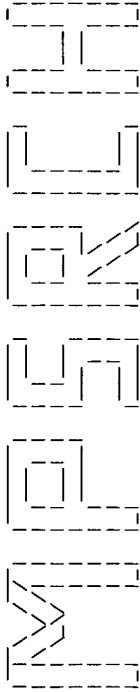


Search completed: Mon Aug 23 13:13:53 1999  
Job time : 10 secs.

---



\*\*\*\*\*



Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:12:32 1999: MasPar time 6.52 Seconds  
Tabular output not generated. 125.775 Million cell updates/sec

Title: >US-09-049-696-46  
Description: (1-29) from US09049696.pep  
Perfect Score: 186  
Sequence: 1 ITVSKTKNKDTSKFPSPLVVYVNIHQAS 29

Scoring table: PAM 150  
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 30.525; Variance 37.677; scale 0.810

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	99	53.2	903	1	ECLC_BOVIN EPITHELIAL CHLORIDE CH	9.47e-09
2	75	40.3	537	1	GUX1_PENJA EXOGLUCANASE I PRECURS	4.16e-03
3	72	38.7	394	1	YEI0_HAEIN HYPOTHETICAL PROTEIN H	1.84e-02
4	69	37.1	976	1	PDR3_YEAST PLEIOTROPIC DRUG RESIS	7.82e-02
5	67	36.0	88	1	PTHP_MYCCA PHOSPHOCARRIER PROTEIN	2.01e-01
6	65	34.9	229	1	YCX2_CABEL HYPOTHETICAL 27.6 KD P	5.05e-01
7	65	34.9	506	1	YCX2_EUGR HYPOTHETICAL 59.8 KD P	5.05e-01
8	64	34.4	462	1	GUNB_FUSOX PUTATIVE ENDOGLUCANASE	7.96e-01
9	64	34.4	960	1	DLG1_DROME LETHAL(1)DISCS LARGE-1	7.96e-01
10	64	34.4	1301	1	YHAG_ECOLI PROBABLE ATP-DEPENDENT	7.96e-01
11	63	33.9	523	1	YHAG_ECOLI HYPOTHETICAL 56.4 KD P	1.25e+00
12	62	33.3	339	1	ENDA_HALVO TRNA-INTRON ENDONUCLEA	1.94e+00
13	62	33.3	449	1	GBA2_YEAST GUANINE NUCLEOTIDE-BIN	1.94e+00
14	62	33.3	780	1	YAD3_SCHPO HYPOTHETICAL 88.2 KD P	1.94e+00
15	61	32.8	1745	1	ZOI_MOUSE TIGHT JUNCTION PROTEIN	3.00e+00
16	60	32.3	148	1	TRBH_RHISN PROBABLE CONJUGAL TRAN	4.63e+00
17	60	32.3	183	1	YGVJ_ECOLI HYPOTHETICAL 20.5 KD P	4.63e+00
18	59	31.7	178	1	Y20K_SSVI HYPOTHETICAL 20.4 KD P	7.08e+00
19	59	31.7	285	1	YDH5_SCHPO HYPOTHETICAL 32.8 KD P	7.08e+00
20	59	31.7	566	1	HEMA_IJAZN HEMAGGLUTININ PRECURSO	7.08e+00
21	59	31.7	567	1	HXTA_YEAST HEXOSE TRANSPORTER HXT	7.08e+00
22	59	31.7	1122	1	YG3C_YEAST HYPOTHETICAL 128.8 KD	7.08e+00
23	59	31.7	3712	1	ACVS_CEPAC DELTA-(L-ALPHA-AMINOAD	7.08e+00

ALIGNMENTS

RESULT	1	ECLC_BOVIN	STANDARD;	PRT;	903 AA.
ID	AC	F54281;			
DT	01-OCT-1996	(REL. 34, CREATED)			
DT	01-OCT-1996	(REL. 34, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)			
DE	EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE CHANNEL).				
DE	OS BOS TAURUS (BOVINE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=TRACHEA;				
RX	MEDLINE: 96125078.				
RA	CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,				
RA	ARRATE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;				
RT	"Cloning of an epithelial chloride channel from bovine trachea.";				
RL	J. BIOL. CHEM. 270:31016-31026(1995).				
CC	!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME; MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND TRANSEPITHELIAL TRANSPORT.				
CC	!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).				
CC	!- TISSUE SPECIFICITY: TRACHEA.				
CC	!- PTM: PHOSPHORYLATED BY CAM-KINASE II.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
DR	EMBL: U36445; G1184056;				
KW	IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL; CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.				
KW	TRANSMEM 7 27 POTENTIAL.				
FT	TRANSMEM 331 351 POTENTIAL.				
FT	TRANSMEM 617 637 POTENTIAL.				
FT	TRANSMEM 883 903 POTENTIAL.				
FT	CARBOHYD 75 75 POTENTIAL.				
FT	CARBOHYD 278 278 POTENTIAL.				
FT	CARBOHYD 360 360 POTENTIAL.				
FT	CARBOHYD 372 372 POTENTIAL.				
FT	CARBOHYD 504 504 POTENTIAL.				

RESULT	4	PDR3_YEAST	STANDARD;	PRT;	976 AA.
ID	AC	P33200;			
DT	DT	01-FEB-1994 (REL. 28, CREATED)			
DT	DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)			
DT	DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	DE	PLEIOTROPIC DRUG RESISTANCE REGULATORY PROTEIN 3.			
GN	GN	PDR3 OR YBL005W OR YBL0323.			
OS	OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).			
OC	OC	EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;			
OC	OC	SACCHAROMYCETACEAE; SACCHAROMYCES.			
RP	RP	[1]			
RN	RN	SEQUENCE FROM N.A.			

```

RC STRAIN=S288C;
RA DELAVEAU T., DELAHODDE A., CARVAJAL E., SUBIK J., JACQ C.;
RL SUBMITTED (JUN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 1-191 FROM N.A.
RX STRAIN=S288C;
RX MEDLINE: 93070613.
RA DELAVEAU T., JACQ C., PEREA J.;
RT "Sequence of a 12.7 kb segment of yeast chromosome II identifies a
RT PDR-like gene and several new open reading frames.";
RL YEAST 8:761-768(1992).
CC -!- FUNCTION: PDR1 AND PDR3 JOINTLY CONTROL THE TRANSCRIPTION LEVEL OF
CC BOTH SN2 AND PDR3.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC CLUSTER DOMAIN.
CC -!- SIMILARITY: STRONG, TO PDR1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X79803; G506478;
DR EMBL: Z35786; G535984;
DR EMBL: S47695; G259052;
DR PIR: S45738; S45738.
DR PIR: S47186; S47186.
DR SGD: L0001363; PDR3.
DR PROSITE: PS00463; ZN2_CYS6_FUNGAL_1; 1.
DR PROSITE: PS00463; ZN2_CYS6_FUNGAL_2; 1.
DR PFAM: PF00172; zn_c1us; 1.
DR HSP: P08657; 1CLD.
DR TRANSCRIPTION REGULATION; ACTIVATOR: DNA-BINDING; NUCLEAR PROTEIN;
KW ZINC; METAL-BINDING.
FT DNA_BIND 15 41 ZN(2)-CYS(6), FUNGAL-TYPE.
SQ SEQUENCE 976 AA; 112570 MW; 0C492188 CRC32;

Query Match 37.1%; Score 69; DB 1; Length 976;
Best Local Similarity 56.3%; Pred. No. 7.82e-02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 127 KDSKYONQLVYQNI 142
QY 9 KDTSKFPLVYVYANI 24

RESULT 5
ID PTHP_MYCCA STANDARD; PRT; 88 AA.
AC P45611;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE PHOSPHOCARRIER PROTEIN HPR (HISTIDINE-CONTAINING PROTEIN).
GN PTHP.
OS MYCOPLASMA CAPRICOLUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC CAPRICOLUM GROUP.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343 / KID;
RX MEDLINE: 94075343.
RA ZHU P.-P., REIZER J., REIZER A., PETERKOFKY A.;
RT "Unique monocistronic operon (ptsh) in Mycoplasma capricolum encoding
RT the phosphocARRIER protein, Hpr, of the phosphoenolpyruvate:sugar
RT phosphotransferase system. Cloning, sequencing, and characterization
RT of ptsh."
RL J. BIOL. CHEM. 268:26531-26540(1993).
RN [2]
RP 3D-STRUCTURE MODELLING.

RX MEDLINE: 96187458.
RA CHURCH W.B., PALMER A., WATHEY J.C., KITSON D.H.;
RT "Homology modeling of histidine-containing phosphocARRIER protein and
RT eosinophil-derived neurotoxin: construction of models and comparison
RT with experiment.";
RL PROTEINS 23:422-430(1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE: 96027100.
RA PIEPER U., KARADIA G., ZHU P.-P., PETERKOFKY A., HERZBERG O.;
RT "Structural evidence for the evolutionary divergence of mycoplasma
RT from gram-positive bacteria: the histidine-containing phosphocARRIER
RT protein."
RL STRUCTURE 3:781-790(1995).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE PHOSPHORYL GROUP FROM PHOSPHOENOLPYRUVATE
CC (PEP) IS TRANSFERRED TO THE PHOSPHORYL CARRIER PROTEIN HPR BY
CC ENZYME I. PHOSPHO-HPR THEN TRANSFERS IT TO THE PERMEASE (ENZYMES
CC II/III). HPR IS COMMON TO ALL PTS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: TO ALL OTHER HPR OR HPR DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L22432; G435097;
DR PDB: 1PCH; 15-OCT-95.
DR PROSITE: PS00369; PTS_HPR_HIS; 1.
DR PROSITE: PS00369; PTS_HPR_SER; 1.
DR PFAM: PF00381; PTS-HPR; 1.
DR PHOSPHOTRANSFERASE SYSTEM; SUGAR TRANSPORT; PHOSPHORYLATION;
KW 3D-STRUCTURE.
FT INIT_MET 0 0
FT MOD_RES 14 14 PHOSPHORYLATION.
SQ SEQUENCE 88 AA; 9287 MW; 0FA866E4 CRC32;

Query Match 36.0%; Score 67; DB 1; Length 88;
Best Local Similarity 47.4%; Pred. No. 2.01e-01;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 23 KEAKFSSNITIANEKQ 41
QY 9 KDTSKFPLVYVYANIRQ 27

RESULT 6
ID YPT3_CAEEL STANDARD; PRT; 229 AA.
AC P41881;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 27.6 KD PROTEIN F37A4.3 IN CHROMOSOME III.
GN F37A4.3.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEAE; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA FULTON L., WATERSTON R.;
RL SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; U00032; G458961; -;  
DR WORMPEP; F37A4.3; CE00709.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 229 AA; 27604 MW; 42AE8C5 CRC32;

Query Match 34.9%; Score 65; DB 1; Length 229;  
Best Local Similarity 44.4%; Pred. No. 5.05e-01;  
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 125 KLSKFDHSLPIYTIIRD 142  
|: ||| | | | | |:  
QY 9 KDTSKFSPLVVYANIRQ 26

RESULT 7  
ID YCX2\_EUGGR STANDARD; PRT; 506 AA.

AC P31921;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 59.8 KD PROTEIN IN PSBD INTRON 8 (ORF506).  
OS EUGLENA GRACILIS.  
OG CHLOROPLAST.  
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=Z;  
RX MEDLINE; 93347989.

RA HALLICK R.B., HONG L., DRAGER R.G., FAVREAU M.R., MONFORT A.,  
ORSAT B., SPIELMANN A., STUTZ E.;  
RT "Complete sequence of Euglena gracilis chloroplast DNA.";  
RL NUCLEIC ACIDS RES. 21:3537-3544(1993).  
CC -!- SIMILARITY: TO GROUP II INTRON MATURASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; Z11874; -; NOT ANNOTATED\_CDS.

DR PIR; S34497; S34497.  
KW CHLOROPLAST; HYPOTHETICAL PROTEIN; ZINC-FINGER.  
FT ZN\_FING 316 339 C2H2-TYPE (POTENTIAL).  
SQ SEQUENCE 506 AA; 59802 MW; D9E714C0 CRC32;

Query Match 34.9%; Score 65; DB 1; Length 506;  
Best Local Similarity 53.8%; Pred. No. 5.05e-01;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 366 KFSKPLVFDNIK 378  
|: ||| | | | | |:  
QY 13 KFPSPLVVYANIR 25

RESULT 8  
ID GUNB\_FUSOX STANDARD; PRT; 462 AA.

AC P46236;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE PUTATIVE ENDOGLUCANASE TYPE B PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-  
DE GLUCANASE) (CELLULOSE).  
OS FUSARIUM OXYSPORIUM.  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;  
OC HYPOCREALES; HYPOCREACEAE; MITOSPORIC HYPOCREACEAE; FUSARIUM.  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95047531.  
RA SHEPPARD P.O., GRANT F.J., OORT P.J., SPRECHER C.A., FOSTER D.C.,  
HAGEN F.S., UPSTALL A., MCKNIGHT G.L., O'HARA P.J.;  
RT "The use of conserved cellulase family-specific sequences to clone  
cellulase homologue cDNAs from Fusarium oxysporum.";  
RL GENE 150:163-167(1994).

CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC  
CC LINKAGES IN CELLULOSE.

CC -!- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).  
CC -!- SIMILARITY: BELONGS TO CELLULOSE FAMILY B (FAMILY 6 OF GLYCOSYL  
CC HYDROLASES).

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; L29377; G520815; -;

DR PROSITE; PS00562; CBD\_FUNGAL; 1.  
DR PROSITE; PS00655; GLYCOSYL\_HYDROL\_F6\_1; 1.  
DR PROSITE; PS00656; GLYCOSYL\_HYDROL\_F6\_2; 1.  
DR PFAM; PF00734; CBD\_fungal; 1.  
DR PFAM; PF01341; Glycosyl\_hydr21; 1.  
DR HSSP; P07987; 1CB2.

DR KW CELLULOSE DEGRADATION; HYDROLASE; GLYCOSIDASE; GLYCOPROTEIN; SIGNAL.  
FT SIGNAL 1 16 POTENTIAL.  
FT CHAIN 17 462 PUTATIVE ENDOGLUCANASE TYPE B.  
FT DOMAIN 17 65 CELLULOSE-BINDING.  
FT DOMAIN 66 99 LINKER.

FT DOMAIN 100 462 CATALYTIC.  
FT ACT\_SITE 190 190 BY SIMILARITY.  
FT ACT\_SITE 236 236 PROTON DONOR (BY SIMILARITY).  
FT ACT\_SITE 416 416 NUCLEOPHILE (BY SIMILARITY).

FT DISULFID 33 50 BY SIMILARITY.  
FT DISULFID 44 60 BY SIMILARITY.  
FT DISULFID 191 250 BY SIMILARITY.  
FT DISULFID 383 430 BY SIMILARITY.  
FT CARBOHYD 37 37 POTENTIAL.  
FT CARBOHYD 223 223 POTENTIAL.  
FT CARBOHYD 272 272 POTENTIAL.  
FT CARBOHYD 317 317 POTENTIAL.  
SQ SEQUENCE 462 AA; 49207 MW; 7B45191F CRC32;

Query Match 34.4%; Score 64; DB 1; Length 462;  
Best Local Similarity 43.8%; Pred. No. 7.96e-01;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 169 KANKAGKYAGQFVYV 184  
|: || | | | | | |:  
QY 6 KTNKDTSKFSPLVVY 21

RESULT 9  
ID DLG1\_DROME STANDARD; PRT; 960 AA.

AC P31007;  
DT 01-JUL-1993 (REL. 26, CREATED)  
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE LETHAL(1)DISCS LARGE-1 TUMOR SUPPRESSOR PROTEIN.  
GN DLG1.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 91330294.  
RA WOODS D.F., BRYANT P.J.;  
RT "The discs-large tumor suppressor gene of Drosophila encodes a



RESULT 14



**This Page Blank (uspto)**



\*\*\*\*\*

WATERMAN

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:13:04 1999; MasPar time 15.31 Seconds  
103.366 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-46  
Description: (1-29) from US09049596.pep  
Perfect Score: 186  
Sequence: 1 ITVTSKTKNDISKFPSPVLYVNIROGAS 29

Scoring table: PAM 150  
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spiremb19  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 29.567; Variance 36.073; scale 0.820

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	143	76.9	913	11	088826 GOR-5 PROTEIN.	6.04e-21
2	94	50.5	794	6	018742 LU-ECAM-1.	1.12e-07
3	94	50.5	820	6	018743 LU-ECAM-1.	1.12e-07
4	94	50.5	905	6	018741 LU-ECAM-1.	1.12e-07
5	80	43.0	901	11	088860 CHLORIDE CHANNEL CACC.	2.73e-04
6	73	39.2	379	1	032705 CITRATE SYNTHASE (EC 4	1.06e-02
7	71	38.2	147	1	028317 CONSERVED HYPOTHETICAL	2.92e-02
8	70	37.6	405	10	049038 PLASMID GCH3937, COMPL	4.80e-02
9	70	37.6	405	10	049026 PLASMID GCH7220, COMPL	4.80e-02
10	66	35.5	421	2	059713 SALICYLATE HYDROXYLASE	3.38e-01
11	63	33.9	419	5	061519 F17E9.8 PROTEIN.	1.38e+00
12	63	33.9	1269	5	043993 PSIA.	1.38e+00
13	63	33.9	2864	14	042975 POLYPROTEIN (FRAGMENT)	1.38e+00
14	63	33.9	2864	14	042974 POLYPROTEIN (FRAGMENT)	1.38e+00
15	63	33.9	2864	14	042973 POLYPROTEIN (FRAGMENT)	1.38e+00
16	63	33.9	2864	14	042976 POLYPROTEIN (FRAGMENT)	1.38e+00
17	63	33.9	3010	14	Q02828 GENOME POLYPROTEIN [CO	1.38e+00
18	63	33.9	3010	14	Q02829 GENOME POLYPROTEIN [CO	1.38e+00
19	63	33.9	3010	14	Q02969 POLYPROTEIN.	1.38e+00
20	63	33.9	3010	14	Q02970 POLYPROTEIN.	1.38e+00

21	63	33.9	3010	14	092972 POLYPROTEIN.	1.38e+00
22	63	33.9	3010	14	092971 POLYPROTEIN.	1.38e+00
23	63	33.9	4436	1	058659 4436AA LONG HYPOTHETIC	1.38e+00
24	62	33.3	447	14	089481 POLYPROTEIN (FRAGMENT)	2.19e+00
25	62	33.3	449	1	027627 PHOSPHOMANNOMUTASE.	2.19e+00
26	62	33.3	1336	5	Q21216 K04D7.5 PROTEIN.	2.19e+00
27	62	33.3	1524	13	091674 POLYPROTEIN (OVOCHYMAS	2.19e+00
28	61	32.8	321	10	039062 CP33 PRECURSOR (FRAGME	3.45e+00
29	61	32.8	329	10	039061 CP33.	3.45e+00
30	61	32.8	453	9	080281 BAN PROTEIN.	3.45e+00
31	61	32.8	480	4	043437 RHO-TYPE GTPASE-ACTIVA	3.45e+00
32	61	32.8	587	4	043182 RHO-TYPE GTPASE-ACTIVA	3.45e+00
33	61	32.8	1371	5	001777 SIMILAR TO PROTEIN-TYR	3.45e+00
34	61	32.8	3010	14	P88803 POLYPROTEIN.	3.45e+00
35	61	32.8	3010	14	P90191 POLYPROTEIN.	3.45e+00
36	60	32.3	167	1	Q58561 HYPOTHETICAL PROTEIN M	5.41e+00
37	60	32.3	326	3	074553 CONSERVED HYPOTHETICAL	5.41e+00
38	60	32.3	447	14	089169 POLYPROTEIN (FRAGMENT)	5.41e+00
39	60	32.3	665	5	Q23906 G-1 PRODUCT.	5.41e+00
40	60	32.3	712	5	Q23873 G-1 ORF.	5.41e+00
41	60	32.3	2034	5	Q17388 FER-1.	5.41e+00
42	60	32.3	3008	14	Q39929 HCV POLYPROTEIN.	5.41e+00
43	60	32.3	3010	14	Q68788 HCV POLYPROTEIN.	5.41e+00
44	60	32.3	3010	14	Q81757 POLYPROTEIN.	5.41e+00
45	59	31.7	838	3	074712 HISTIDINOL DEHYDROGENA	8.41e+00

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	913 AA.
ID	088826	PRELIMINARY;	PRT;	913 AA.
AC	088826;			
DT	01-NOV-1998	(TREMBLREL. 08, CREATED)		
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)		
DE	GOR-5 PROTEIN.			
GN	GOR-5.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			
OC	SCIUROGNATHI; MURIDAE; MURINAE; MUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=INTESTINE;			
RA	KOMIYA T., TANIGAWA Y., HIROHASHI S.;			
RT	"Cloning of a gene, gob-5, which is expressed in intestinal goblet			
RT	cells in mice."			
RL	SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL: AB017156; D1034712.			
SQ	SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;			

Query Match 76.9%; Score 143; DB 11; Length 913;  
Best Local Similarity 72.4%; Pred. No. 6.04e-21;  
Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db	592	ITVTPVVKNTGKFPSPVTVYASIRQAS	620
QY	1	ITVTSKTKNDISKFPSPVLYVNIROGAS	29
		:   :       :    :	

RESULT	2	PRELIMINARY;	PRT;	794 AA.
ID	018742	PRELIMINARY;	PRT;	794 AA.
AC	018742;			

DT	01-JAN-1998	(TREMBLREL. 05, CREATED)
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT	01-AUG-1998	(TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE	LU-ECAM-1.	
OS	BOS TAURUS (BOVINE).	
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;	
OC	ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LUNG;	
RA	EUBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,	

RA GOODWIN A., PAULI B.U.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF001262; G2623765; -  
 SQ SEQUENCE 794 AA; 88509 MW; 2D450EEB CRC32;

Query Match 50.5%; Score 94; DB 6; Length 794;  
 Best Local Similarity 37.0%; Pred. No. 1.12e-07;  
 Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAMHSQHTAHYPSMIVYAVSQS 625

QY 1 ITVTSKNTKDTSKFSPPLVYVYANIQQ 27

RESULT 3

ID O18743 PRELIMINARY; PRT: 820 AA.

AC O18743;

DT 01-JAN-1998 (TREMREL. 05, CREATED)

DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)

DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)

DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVINA; BOVINE; BOS.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;

RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,

RA GOODWIN A., PAULI B.U.;

RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF001263; G2623767; -

SQ SEQUENCE 820 AA; 91464 MW; 171ED21C CRC32;

Query Match 50.5%; Score 94; DB 6; Length 820;

Best Local Similarity 37.0%; Pred. No. 1.12e-07;

Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAMHSQHTAHYPSMIVYAVSQS 625

QY 1 ITVTSKNTKDTSKFSPPLVYVYANIQQ 27

RESULT 4

ID O18741 PRELIMINARY; PRT: 905 AA.

AC O18741;

DT 01-JAN-1998 (TREMREL. 05, CREATED)

DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVINA; BOVINE; BOS.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;

RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,

RA GOODWIN A., PAULI B.U.;

RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF001261; G2623763; -

DR PFAM: PF00041; fn3; 1

SQ SEQUENCE 905 AA; 101005 MW; 86E65459 CRC32;

Query Match 50.5%; Score 94; DB 6; Length 905;

Best Local Similarity 37.0%; Pred. No. 1.12e-07;

Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 599 VIATAMHSQHTAHYPSMIVYAVSQS 625

QY 1 ITVTSKNTKDTSKFSPPLVYVYANIQQ 27

RESULT 5

ID O88860 PRELIMINARY; PRT: 901 AA.

AC O88860;  
 DT 01-NOV-1998 (TREMREL. 08, CREATED)  
 DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE CHLORIDE CHANNEL CACC.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
 OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]

RP SEQUENCE FROM N.A.

RA ROMO L., MUSANTE L., CINTI R., MORAN O., SERI M., GALIETTA L.J.V.;

RT "Characterization of a murine gene homologous to the bovine CaCC

RT chloride channel";

RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF052746; G3560547; -

SQ SEQUENCE 901 AA; 100039 MW; E52BF02E CRC32;

Query Match 43.0%; Score 80; DB 11; Length 901;

Best Local Similarity 37.5%; Pred. No. 2.73e-04;

Matches 9; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 599 TAHMSQTAQYPSRMIVYAVSQS 622

QY 4 TSKNTKDTSKFSPPLVYVYANIQQ 27

RESULT 6

ID O32705 PRELIMINARY; PRT: 379 AA.

AC O32705;

DT 01-JAN-1998 (TREMREL. 05, CREATED)

DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE CITRATE SYNTHASE (EC 4.1.3.7) (FRAGMENT).

OS HALOBACTERIUM VOLCANII (HALOFERAX VOLCANII).

OC ARCHAEA; EURYARCHAEOTA; HALOBACTERIALES; HALOBACTERIACEAE; HALOFERAX.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-WFD II;

RA MADDOCKS D.G., CONNARIS H., HOUGH D.W., DANSON M.J.;

RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AJ002075; E1154178; -

DR PFAM: PF00285; citrate\_synt; 1.

KW LYASE

FT NON\_TER 379 379

SQ SEQUENCE 379 AA; 41848 MW; 480EAC2F CRC32;

Query Match 39.2%; Score 73; DB 1; Length 379;

Best Local Similarity 45.5%; Pred. No. 1.06e-02;

Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 126 KKRITAKMPSVLAAYAFRRG 147

QY 6 KTKNTKDTSKFSPPLVYVYANIQQ 27

RESULT 7

ID O28317 PRELIMINARY; PRT: 147 AA.

AC O28317;

DT 01-JAN-1998 (TREMREL. 05, CREATED)

DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)

DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)

DE CONSERVED HYPOTHETICAL PROTEIN.

GN AF1962.

OS ARCHAEoglobus FULGIDUS.

OC ARCHAEA; EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobACEAE;

OC ARCHAEoglobus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;

RX MEDLINE: 98049343.

RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,

RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,

RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,

RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,  
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,  
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,  
 RA OVERBEER R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,  
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYKES S.M.,  
 RA SADON P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,  
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,  
 RA VENTER J.C.;  
 RT "The complete genome sequence of the hyperthermophilic,  
 RT sulphate-reducing archaeon *Archaeoglobus fulgidus*.";  
 RL NATURE 390:364-370(1997).  
 DR EMBL: AF000967; G2648577; -  
 DR TIGR: AF1962; -  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 147 AA; 16481 MW; 11A7C55E CRC32;  
 Query Match 38.2%; Score 71; DB 1; Length 147;  
 Best Local Similarity 47.1%; Pred. No. 2.92e-02;  
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Db 75 SREPERLIVYSADRHNA 91  
 QY 12 SKFSPPLVYVYANIROGA 28  
 RESULT 8  
 ID O49038; PRELIMINARY; PRT; 405 AA.  
 AC O49038;  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE PLASMID GCH3937, COMPLETE SEQUENCE.  
 OS GRACILARIA CHILENSIS.  
 OG PLASMID GCH3937.  
 OC EUKARYOTA; RHODOPHYTA; FLORIDOPHYCEAE; GRACILARIALES; GRACILARIACEAE;  
 OC GRACILARIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GOFF L.J., MOON D.A.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF034721; G2921378; -  
 KW PLASMID.  
 SQ SEQUENCE 405 AA; 47292 MW; 371FF41E CRC32;  
 Query Match 37.6%; Score 70; DB 10; Length 405;  
 Best Local Similarity 33.3%; Pred. No. 4.80e-02;  
 Matches 8; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
 Db 249 ITMSSKSNHEVCEFGSLLSFNRI 272  
 QY 1 ITVTSKTKNDTSKFPSPLVVYANI 24  
 RESULT 9  
 ID O49026; PRELIMINARY; PRT; 405 AA.  
 AC O49026;  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE PLASMID GCH7220, COMPLETE SEQUENCE.  
 OS GRACILARIA CHILENSIS.  
 OG PLASMID GCH7220.  
 OC EUKARYOTA; RHODOPHYTA; FLORIDOPHYCEAE; GRACILARIALES; GRACILARIACEAE;  
 OC GRACILARIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA GOFF L.J., MOON D.A.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF034719; G2921358; -  
 KW PLASMID.  
 SQ SEQUENCE 405 AA; 47275 MW; 7A6BF5B0 CRC32;  
 Query Match 37.6%; Score 70; DB 10; Length 405;

Best Local Similarity 33.3%; Pred. No. 4.80e-02;  
 Matches 8; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
 Db 249 ITMSSKSNHEVCEFGSLLSFNRI 272  
 QY 1 ITVTSKTKNDTSKFPSPLVVYANI 24  
 RESULT 10  
 ID Q59713; PRELIMINARY; PRT; 421 AA.  
 AC Q59713;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE SALICYLATE HYDROXYLASE (EC 1.14.13.1) (SALICYLATE 1-MONOOXYGENASE)  
 DE (FRAGMENT).  
 OS PSEUDOMONAS PUTIDA.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;  
 OC PSEUDOMONAS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MIZUGUCHI M., OHNISHI K., SUZUKI K., SATO H., NAKAMURA Y., ITAGAKI E.,  
 RA FUKUDA S.;  
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AB010714; D1025490; -  
 KW OXIDOREDUCTASE.  
 SQ SEQUENCE 421 AA; 45258 MW; D992F4F8 CRC32;  
 Query Match 35.5%; Score 66; DB 2; Length 421;  
 Best Local Similarity 38.9%; Pred. No. 3.38e-01;  
 Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
 Db 345 EASDIPALLEVYDVERG 362  
 QY 10 DTSKFPSPLVVYANIRQG 27  
 RESULT 11  
 ID O61519; PRELIMINARY; PRT; 419 AA.  
 AC O61519;  
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE F17E9.8 PROTEIN.  
 GN F17E9.8.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNITEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE: 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL NATURE 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WOESSNER J.;  
 RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;

```
RA WATERSTON R.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AF047656; G2935379; -.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
KW POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; TRANSMEMBRANE.
SQ SEQUENCE 419 AA; 48226 MW; 8B4808A1 CRC32;

Query Match      33.9%; Score 63; DB 5; Length 419;
Best Local Similarity 41.7%; Pred. No. 1.38e+00;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Db 383 IVTAETNEEINFFNSLVGYII 406
|::: |::: |::: |::: |::: |
QY 1 ITVTSKTKNDTSKFPSPLVVYANI 24

RESULT 12
ID O43993 PRELIMINARY; PRT: 1269 AA.
AC O43993;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE PSLA.
DE PSLA.
GN PSLA.
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KAX3;
RA YASUKAWA H., MOHANTY S., FIRTEL R.A.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF038919; G2766696; -.
SQ SEQUENCE 1269 AA; 148589 MW; 46033229 CRC32;

Query Match      33.9%; Score 63; DB 5; Length 1269;
Best Local Similarity 34.8%; Pred. No. 1.38e+00;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Db 399 TSNKKDSIVKFKDITIIYLNVRQ 421
|::: |::: |::: |::: |::: |
QY 4 TSKTKNDTSKFPSPLVVYANI 26

RESULT 13
ID O92975 PRELIMINARY; PRT: 2864 AA.
AC O92975;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE POLYPROTEIN (FRAGMENT).
OS HEPATITIS C VIRUS (HCV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;
OC HEPATITIS C-LIKE VIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HC-J4;
RX MEDLINE; 98240944.
RA YANAGI M., ST CLARE M., SHAPIRO M., EMERSON S.U., PURCELL R.H.,
RA BUKH J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL VIROLOGY 244:161-172(1998).
DR EMBL: AF054258; G3098653; -.
KW POLYPROTEIN.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310415 MW; DF789764 CRC32;

Query Match      33.9%; Score 63; DB 14; Length 2864;
Best Local Similarity 36.4%; Pred. No. 1.38e+00;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

RA WATERSTON R.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AF047656; G2935379; -.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
KW POSTSYNAPTIC MEMBRANE; IONIC CHANNEL; GLYCOPROTEIN; TRANSMEMBRANE.
SQ SEQUENCE 419 AA; 48226 MW; 8B4808A1 CRC32;

Query Match      33.9%; Score 63; DB 5; Length 419;
Best Local Similarity 41.7%; Pred. No. 1.38e+00;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Db 383 IVTAETNEEINFFNSLVGYII 406
|::: |::: |::: |::: |::: |
QY 1 ITVTSKTKNDTSKFPSPLVVYANI 24

RESULT 12
ID O43993 PRELIMINARY; PRT: 1269 AA.
AC O43993;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE PSLA.
DE PSLA.
GN PSLA.
OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KAX3;
RA YASUKAWA H., MOHANTY S., FIRTEL R.A.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF038919; G2766696; -.
SQ SEQUENCE 1269 AA; 148589 MW; 46033229 CRC32;

Query Match      33.9%; Score 63; DB 5; Length 1269;
Best Local Similarity 34.8%; Pred. No. 1.38e+00;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Db 399 TSNKKDSIVKFKDITIIYLNVRQ 421
|::: |::: |::: |::: |::: |
QY 4 TSKTKNDTSKFPSPLVVYANI 26

RESULT 13
ID O92975 PRELIMINARY; PRT: 2864 AA.
AC O92975;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE POLYPROTEIN (FRAGMENT).
OS HEPATITIS C VIRUS (HCV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;
OC HEPATITIS C-LIKE VIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HC-J4;
RX MEDLINE; 98240944.
RA YANAGI M., ST CLARE M., SHAPIRO M., EMERSON S.U., PURCELL R.H.,
RA BUKH J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL VIROLOGY 244:161-172(1998).
DR EMBL: AF054258; G3098653; -.
KW POLYPROTEIN.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310415 MW; DF789764 CRC32;

Query Match      33.9%; Score 63; DB 14; Length 2864;
Best Local Similarity 36.4%; Pred. No. 1.38e+00;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 2268 ISVAAEILRKSRKFPSPALPIWA 2289
|::: |::: |::: |::: |::: |
QY 1 ITVTSKTKNDTSKFPSPLVVYA 22

RESULT 14
ID O92974 PRELIMINARY; PRT: 2864 AA.
AC O92974;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE POLYPROTEIN (FRAGMENT).
OS HEPATITIS C VIRUS (HCV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;
OC HEPATITIS C-LIKE VIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HC-J4;
RX MEDLINE; 98240944.
RA YANAGI M., ST CLARE M., SHAPIRO M., EMERSON S.U., PURCELL R.H.,
RA BUKH J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL VIROLOGY 244:161-172(1998).
DR EMBL: AF054257; G3098651; -.
KW POLYPROTEIN.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310481 MW; E169861C CRC32;

Query Match      33.9%; Score 63; DB 14; Length 2864;
Best Local Similarity 36.4%; Pred. No. 1.38e+00;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 2268 ISVAAEILRKSRKFPSPALPIWA 2289
|::: |::: |::: |::: |::: |
QY 1 ITVTSKTKNDTSKFPSPLVVYA 22

RESULT 15
ID O92973 PRELIMINARY; PRT: 2864 AA.
AC O92973;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE POLYPROTEIN (FRAGMENT).
OS HEPATITIS C VIRUS (HCV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FLAVIVIRIDAE;
OC HEPATITIS C-LIKE VIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HC-J4;
RX MEDLINE; 98240944.
RA YANAGI M., ST CLARE M., SHAPIRO M., EMERSON S.U., PURCELL R.H.,
RA BUKH J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL VIROLOGY 244:161-172(1998).
DR EMBL: AF054253; G3098644; -.
KW POLYPROTEIN.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310456 MW; 35A5D7A3 CRC32;

Query Match      33.9%; Score 63; DB 14; Length 2864;
Best Local Similarity 36.4%; Pred. No. 1.38e+00;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 2268 ISVAAEILRKSRKFPSPALPIWA 2289
|::: |::: |::: |::: |::: |
QY 1 ITVTSKTKNDTSKFPSPLVVYA 22

Search completed: Mon Aug 23 13:13:25 1999
Job time : 21 secs.
```

\*\*\*\*\*

WQEQELH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:59:12 1999; MasPar time 9.93 Seconds  
Tabular output not generated. 62.105 Million cell updates/sec

Title: >US-09-049-696-46  
Description: (1-29) from US09049696.pep  
Perfect Score: 29  
Sequence: 1 ITVTSKTNKDTSKPPSLVYVNIROGAS 29

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 2.007; Variance 0.642; scale 3.126

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	20.7	784	30	W53570	Cucumber raffinose sy
2	5	17.2	9	19	W07127	Synthetic peptide use
3	5	17.2	10	5	R03909	HIV-antibody reactive
4	5	17.2	12	17	R90327	IL-8 epitope 8-24
5	5	17.2	20	24	W24101	Canine immunoglobulin E
6	5	17.2	20	24	W24100	Canine immunoglobulin E
7	5	17.2	26	17	R52494	Human heavy chain sur
8	5	17.2	26	17	R52435	Mouse heavy chain sur
9	5	17.2	32	12	R72117	Mouse anti-human il-6
10	5	17.2	43	2	R11409	Hepatitis B surface a
11	5	17.2	70	2	R11482	Membrane anchoring pe
12	5	17.2	71	2	R11480	Membrane anchoring pe
13	5	17.2	71	9	R47919	Segment of human gamm
14	5	17.2	71	9	R47921	Segment of human gamm
15	5	17.2	79	1	P91702	Protein increasing pu
16	5	17.2	102	2	P70146	Product of the second

109	24	W15559	TGF beta-1/TGF beta-2	4.15e+02
109	29	W47416	Light chain variable	4.15e+02
111	12	R6304	Human immunoglobulin	4.15e+02
113	27	W23025	Humanized 5D12 monoclonal	4.15e+02
113	27	W23023	Murine 5D12 monoclonal	4.15e+02
116	31	W53994	Anti-CD4 antibody VH2	4.15e+02
116	19	W05823	Humanized ID10 antibody	4.15e+02
117	31	W58504	Human heavy chain sub	4.15e+02
119	12	R66318	Human immunoglobulin	4.15e+02
119	12	R66299	Human immunoglobulin	4.15e+02
120	13	R77303	Variable heavy chain	4.15e+02
120	16	R92086	Human Cos VH region	4.15e+02
121	25	W27552	Human Ab heavy chain	4.15e+02
124	24	W24097	Partial canine immuno	4.15e+02
126	21	W11635	Human anti-RSV monoclonal	4.15e+02
128	16	R92087	Human Cess RV region	4.15e+02
128	10	R54798	SP4-reactive IgM heavy	4.15e+02
130	31	W53993	Human Anti-CD4 antibody	4.15e+02
133	23	W10538	Humanized murine anti	4.15e+02
133	23	W10540	Humanized murine anti	4.15e+02
133	8	R42800	Bi3/B14 HuVH	4.15e+02
137	29	W30273	Heavy chain of HuM4TS	4.15e+02
138	20	W01146	MAB 10.1 heavy chain	4.15e+02
139	13	R67658	Anti-human IL-6 chimera	4.15e+02
139	7	R38315	Sequence of the VH of	4.15e+02
139	13	R67659	Anti-human IL-6 chimera	4.15e+02
142	2	W81064	Sequence of rhinovirus	4.15e+02
148	27	W26239	Monoclonal antibody 1	4.15e+02
151	17	R95910	Opacity associated pr	4.15e+02
163	7	R34220	Sequence of R08/26 sa	4.15e+02
168	31	W28343	Amino acid sequence o	4.15e+02
171	22	W20323	H. pylori cell envelo	4.15e+02
191	33	W48897	Candida albicans CaCd	4.15e+02
191	28	W30380	Candida CaCd42 prote	4.15e+02
191	28	W33897	Candida CaCd42 prote	4.15e+02
206	36	W80412	A secreted protein en	4.15e+02
206	22	W20889	H. pylori transporter	4.15e+02
222	37	W75058	Human secreted protei	4.15e+02
226	4	P30096	Sequence of hepatitis	4.15e+02
226	26	P30060	Surface antigen of HB	4.15e+02
226	1	P80417	Sequence of S protein	4.15e+02
226	3	P60223	Sequence of adv type h	4.15e+02
227	39	W73576	Human nucleic acid bi	4.15e+02
256	2	P70404	ORF 1 gene product of	4.15e+02
264	37	W80482	Islet cell antibody a	4.15e+02
264	2	R05464	Derived protein from	4.15e+02
273	19	W05827	Humanized ID10 antibo	4.15e+02
281	3	P60164	Subtype adv hepatitis	4.15e+02
281	3	P60561	Hepatitis B virus P31	4.15e+02
283	14	R73947	T.sergenti 33 kDa pro	4.15e+02
283	9	R45456	Theileria buffeli mai	4.15e+02
283	9	R45455	Theileria sergenti ma	4.15e+02
285	1	R03622	Watermelon mosaic vir	4.15e+02
288	1	P90460	Alpha-factor profibro	4.15e+02
315	28	W37462	Human Ret ligand RetL	4.15e+02
332	31	W47831	Human melanocortin-4	4.15e+02
332	29	W42378	Homo sapiens mutant m	4.15e+02
332	29	W42377	Homo sapiens mutant m	4.15e+02
332	36	W79687	Melanocortin-4 recept	4.15e+02
332	28	W37324	Human melanocortin-4	4.15e+02
332	29	W42379	Homo sapiens mutant m	4.15e+02
332	22	W19704	Melanocortin-4 recept	4.15e+02
332	16	R48732	G-protein coupled hum	4.15e+02
332	19	W02704	G-protein coupled hum	4.15e+02
336	18	W00860	Human T-lymphocyte TL	4.15e+02
336	36	W79094	Human secreted protei	4.15e+02
336	36	W18051	DNAX accessory molecu	4.15e+02
336	23	R20812	T-lymphocyte TlISA An	4.15e+02
339	22	W20200	H. pylori transmembr	4.15e+02
340	34	W61615	Human apoptosis-assoc	4.15e+02
342	37	W75049	Fragment of human sec	4.15e+02
346	28	W37465	Mouse Ret ligand retL	4.15e+02
351	10	R51687	MVP-5180/91 fragment	4.15e+02

90	5	17.2	362 22	W20626	H. pylori transmembra	4.15e+02	163	578 39	W85604	A hexosaminidase enzy	4.15e+02
91	5	17.2	369 18	W02580	HIV-2 transmembrane p	4.15e+02	164	579 7	R34446	N.meningitidis 2394 T	4.15e+02
92	5	17.2	369 29	W1576	Arabidopsis C-14 ster	4.15e+02	165	588 9	R47257	Pre-pro-BPP.	4.15e+02
93	5	17.2	372 1	R05098	Synthetic HIV-2 trans	4.15e+02	166	590 24	W24577	Merzoite apical-end	4.15e+02
94	5	17.2	372 18	W02577	Recombinant HIV-2 tra	4.15e+02	167	590 2	R07505	Fragment of Plasmodiu	4.15e+02
95	5	17.2	376 36	W2189	HSV-2 strain SB5 Cont	4.15e+02	168	616 13	R73005	Aminopectidase O12 cl	4.15e+02
96	5	17.2	376 35	W2044	HSV-2 strain SB5 Cont	4.15e+02	169	616 13	R73007	Aminopectidase O12 cl	4.15e+02
97	5	17.2	376 36	W2121	HSV-2 strain SB5 Cont	4.15e+02	170	618 38	W73312	Cellulose synthase Pc	4.15e+02
98	5	17.2	378 39	W84185	Glial cell line-deriv	4.15e+02	171	622 38	W73311	Cellulose synthase.	4.15e+02
99	5	17.2	379 2	R05110	Pela Pectin lyase rec	4.15e+02	172	624 36	W70494	Human disease related	4.15e+02
100	5	17.2	392 23	W8385	S. typhimurium prgh 9	4.15e+02	173	625 39	W67862	Human secreted protei	4.15e+02
101	5	17.2	392 12	R0740	Prgh gene product.	4.15e+02	174	626 15	R71034	N. crassa glucoamylas	4.15e+02
102	5	17.2	397 39	W84182	A GDNFR-alpha-related	4.15e+02	175	629 30	W33816	Arabidopsis cellulose	4.15e+02
103	5	17.2	397 28	W37461	Mouse Ret ligand RetL	4.15e+02	176	631 19	W00604	Murine myosin VII pro	4.15e+02
104	5	17.2	400 39	W84180	A GDNFR-alpha-related	4.15e+02	177	640 32	W48855	Luciferase-lysoastaphi	4.15e+02
105	5	17.2	400 39	W84186	Glial cell line-deriv	4.15e+02	178	670 18	W02582	HIV-1 gag/HIV-2 TMP f	4.15e+02
106	5	17.2	400 32	W5116	Human GDNF alpha-3 re	4.15e+02	179	670 1	W02599	PSD307.PEP HIV-1 gag/	4.15e+02
107	5	17.2	400 32	W5117	Human GDNF alpha-3 re	4.15e+02	180	682 11	R49144	Product of alternativ	4.15e+02
108	5	17.2	400 28	W37463	Human Ret ligand RetL	4.15e+02	181	682 17	R87154	Alternatively spliced	4.15e+02
109	5	17.2	405 1	P90458	Snake venom fibrolase	4.15e+02	182	684 21	W17591	Thermoanaerobacter CG	4.15e+02
110	5	17.2	405 2	R04082	Gene product of south	4.15e+02	183	684 21	W17583	Thermoanaerobacter CG	4.15e+02
111	5	17.2	406 25	W26604	Senna alpha-galactosi	4.15e+02	184	684 21	W17587	Thermoanaerobacter CG	4.15e+02
112	5	17.2	407 13	R5494	Marek's disease virus	4.15e+02	185	685 2	R11331	Human luteinising hor	4.15e+02
113	5	17.2	410 37	W39335	Human secreted protei	4.15e+02	186	697 19	W00603	Human non-conventiona	4.15e+02
114	5	17.2	416 27	W32113	Human oestrogen recep	4.15e+02	187	699 4	R22187	Sequence of luteinisi	4.15e+02
115	5	17.2	417 26	W23067	Canine IgE heavy chai	4.15e+02	188	708 38	W83394	Human protein-coupled	4.15e+02
116	5	17.2	418 28	W33214	Human oestrogen recep	4.15e+02	189	731 32	W44849	S. pneumoniae penicil	4.15e+02
117	5	17.2	425 8	R33132	VPL/P2A protein.	4.15e+02	190	731 20	W03180	Mouse poly-immunoglob	4.15e+02
118	5	17.2	426 22	W09821	Mouse interleukin-12	4.15e+02	191	793 13	R70234	P. faiciparum E31a.	4.15e+02
119	5	17.2	426 17	R97753	Canine IgE.	4.15e+02	192	804 23	W22709	Leucyl-tRNA synthetas	4.15e+02
120	5	17.2	446 19	W05829	Humanized LD10 antibo	4.15e+02	193	822 33	W69847	Amino acid sequence o	4.15e+02
121	5	17.2	448 30	W53463	Human gp49 HM18 poly	4.15e+02	194	824 14	W00364	Human CDC27.	4.15e+02
122	5	17.2	448 37	W82551	Human LIR-pbm2 protei	4.15e+02	195	824 14	R75848	H-NUC retinoblastoma	4.15e+02
123	5	17.2	449 23	W82210	Cellulomonas fimi end	4.15e+02	196	836 11	R58912	Product of alternativ	4.15e+02
124	5	17.2	453 9	R47263	Pre-pro BMP5.	4.15e+02	197	836 17	R87153	Alternatively spliced	4.15e+02
125	5	17.2	454 2	R07310	Human Bone Morphogeni	4.15e+02	198	852 2	R04029	Env gene product of c	4.15e+02
126	5	17.2	454 7	R36736	Human BMP-5.	4.15e+02	199	856 1	R80803	Sequence of env prote	4.15e+02
127	5	17.2	454 4	R32949	Encodes C-terminal po	4.15e+02	200	857 23	W13052	HIV-2 provirus-encode	4.15e+02
128	5	17.2	458 1	P90463	Alpha-factor profibro	4.15e+02	201	858 1	R80806	Sequence of env prote	4.15e+02
129	5	17.2	460 1	P90462	Alpha-facto profibro	4.15e+02	202	858 4	R20604	Sequence encoded by o	4.15e+02
130	5	17.2	472 34	W69234	FCR-IV protein sequen	4.15e+02	203	858 1	R81779	ROD HIV-2 Env protein	4.15e+02
131	5	17.2	475 21	W1641	Human anti-RSV monocl	4.15e+02	204	877 20	W07205	HIV-1 group O strain	4.15e+02
132	5	17.2	476 21	W01818	Primates anti-human	4.15e+02	205	878 22	W20966	H. pylori flagella-as	4.15e+02
133	5	17.2	476 33	W63761	Macaque primatized 7C	4.15e+02	206	879 37	W72979	Bovine beta-mannosida	4.15e+02
134	5	17.2	477 28	W33212	Human oestrogen recep	4.15e+02	207	879 22	W14265	Bovine beta-mannosida	4.15e+02
135	5	17.2	478 1	P90461	Alpha-factor profibro	4.15e+02	208	884 18	W00160	MEI2 protein kinase P	4.15e+02
136	5	17.2	478 2	R04083	Profibrolase product	4.15e+02	209	887 13	R65492	Pyruvate dehydrogenas	4.15e+02
137	5	17.2	478 1	P90459	Alpha-factor profibro	4.15e+02	210	891 35	W70293	Simian herpesvirus B	4.15e+02
138	5	17.2	485 28	W39422	ACC synthase protein.	4.15e+02	211	891 16	R92746	B virus gB glycoprote	4.15e+02
139	5	17.2	485 29	W47314	Tomato ACC synthase L	4.15e+02	212	891 1	R82677	ENVRN sequence from H	4.15e+02
140	5	17.2	485 29	W47313	Tomato ACC synthase.	4.15e+02	213	904 17	R87147	Protocadherin clone 4	4.15e+02
141	5	17.2	485 3	R15506	Tomato ACC synthase e	4.15e+02	214	904 8	R41757	Nitrate reductase Nia	4.15e+02
142	5	17.2	485 21	W14724	Human oestrogen recep	4.15e+02	215	904 11	R58907	Human protocadherin-4	4.15e+02
143	5	17.2	485 3	R15507	Tomato ACC synthase e	4.15e+02	216	911 23	R10333	Deduced sequence of t	4.15e+02
144	5	17.2	497 39	W89786	Staphylococcus aureus	4.15e+02	217	921 23	W22480	Plasmodium E31a.	4.15e+02
145	5	17.2	498 15	R82424	Capsanthin-capsorubin	4.15e+02	218	932 38	W85040	PKG-green fluorescent	4.15e+02
146	5	17.2	503 27	W27133	Squalene epoxidase pr	4.15e+02	219	933 38	W85039	Green fluorescent pro	4.15e+02
147	5	17.2	504 38	W86245	Mouse ActRIIB4 recep	4.15e+02	220	934 5	P20016	Sequence of p20, VP4,	4.15e+02
148	5	17.2	510 6	R29582	Xenopus activin recep	4.15e+02	221	972 8	R51281	Helminth aminopeptida	4.15e+02
149	5	17.2	512 38	W86243	Mouse ActRIIB2 recep	4.15e+02	222	971 16	R80096	Black widow spider de	4.15e+02
150	5	17.2	521 27	R32303	Arabidopsis thaliana	4.15e+02	223	1007 33	W61246	Streptococcus pneumon	4.15e+02
151	5	17.2	527 15	R90725	Laccase-LCC5.	4.15e+02	224	1018 2	R82115	Fibronectin binding p	4.15e+02
152	5	17.2	528 38	W86244	Mouse ActRIIB3 recep	4.15e+02	225	1027 39	W89806	Staphylococcus aureus	4.15e+02
153	5	17.2	530 28	W33215	Human oestrogen recep	4.15e+02	226	1037 39	W67843	A serine/threonine pr	4.15e+02
154	5	17.2	534 30	W46748	Arabidopsis thaliana	4.15e+02	227	1065 30	W33819	Arabidopsis cellulose	4.15e+02
155	5	17.2	534 26	W36137	A. thaliana inorganic	4.15e+02	228	1081 30	W33817	Arabidopsis cellulose	4.15e+02
156	5	17.2	534 27	W32302	Arabidopsis thaliana	4.15e+02	229	1081 30	W33820	Arabidopsis cellulose	4.15e+02
157	5	17.2	536 38	W86242	Mouse ActRIIB1 recep	4.15e+02	230	1084 30	W33818	Arabidopsis cellulose	4.15e+02
158	5	17.2	536 35	W05406	Mouse cortactin prote	4.15e+02	231	1134 29	W46859	Bacillus thuringiens	4.15e+02
159	5	17.2	548 37	W81459	C. acidovorans polyur	4.15e+02	232	1180 34	W61273	Homo sapiens protocad	4.15e+02
160	5	17.2	572 35	W68487	Mouse ULIP-3 protein.	4.15e+02	233	1214 16	R80097	Black widow spider de	4.15e+02
161	5	17.2	572 35	W68486	Mouse ULIP-1 protein.	4.15e+02	234	1248 35	W69594	Filamentous haemagglu	4.15e+02
162	5	17.2	578 21	W01639	T. harzianum exochiti	4.15e+02	235	1294 36	W75798	Arabidopsis ethylene	4.15e+02

236	5	17.2	1321	17	R88423	Arabisopsis thaliana	4.15e+02	309	4	13.8	35	4	R20831	PND EE293-3.	3.23e+03
237	5	17.2	1592	6	R32925	Glucosyltransferase	4.15e+02	310	4	13.8	35	4	R20950	PND EE667-3.	3.23e+03
238	5	17.2	1805	11	R60126	Rat nestin protein is	4.15e+02	311	4	13.8	35	4	R20835	PND EE297-2.	3.23e+03
239	5	17.2	1805	5	R27204	Rat nestin.	4.15e+02	312	4	13.8	35	3	R14330	HIV-1 amplifier pepti	3.23e+03
240	5	17.2	1972	20	W00024	Smooth muscle myosin	4.15e+02	313	4	13.8	36	39	W8515	Amphotropic hypervari	3.23e+03
241	5	17.2	2150	3	P60045	Sequence of viral pro	4.15e+02	314	4	13.8	36	30	W53053	Tbpl antigenic peptid	3.23e+03
242	5	17.2	2150	2	P81048	Sequence encoded by R	4.15e+02	315	4	13.8	37	13	R65389	Universal immunostimu	3.23e+03
243	5	17.2	2150	2	R05127	Complete human rhinov	4.15e+02	316	4	13.8	37	13	R65383	Universal immunostimu	3.23e+03
244	5	17.2	2185	2	R12141	Enteroviral polypepti	4.15e+02	317	4	13.8	37	19	R98149	Thermostable inorgani	3.23e+03
245	5	17.2	2322	22	W15566	Melanoma-associated c	4.15e+02	318	4	13.8	37	13	R66393	SREBP-1a C-terminal f	3.23e+03
246	5	17.2	2353	19	R99393	Haemophilus adhesio	4.15e+02	319	4	13.8	39	38	W85095	Thyroid hormone recep	3.23e+03
247	5	17.2	3077	1	P93283	Sequence of clone HIV	4.15e+02	320	4	13.8	39	13	R65390	Universal immunostimu	3.23e+03
248	5	17.2	3211	1	P81769	Sequence encoded by t	4.15e+02	321	4	13.8	39	20	W01478	Agammaglobulinemia t	3.23e+03
249	5	17.2	3647	1	R05041	Filamentous haemagglu	4.15e+02	322	4	13.8	39	10	R53537	Thyroid hormone recep	3.23e+03
250	5	17.2	4544	9	R47861	Alpha 2-Macroglobulin	4.15e+02	323	4	13.8	41	39	W88514	Amphotropic hypervari	3.23e+03
251	5	17.2	4544	9	R60517	Human alpha-2-MR.	4.15e+02	324	4	13.8	41	2	P81890	Sequence encoded by h	3.23e+03
252	4	13.8	7	17	R90422	Antibody 561 displaci	3.23e+03	325	4	13.8	43	39	W91061	Flanking sequence use	3.23e+03
253	4	13.8	7	3	R13525	Analgesic heptapeptid	3.23e+03	326	4	13.8	43	39	W91059	Flanking sequence use	3.23e+03
254	4	13.8	9	28	W49517	Human leucocyte anti	3.23e+03	327	4	13.8	43	39	W93474	PCANAB vector encodi	3.23e+03
255	4	13.8	10	34	W65749	Fibronectin binding p	3.23e+03	328	4	13.8	43	39	W91057	Flanking sequence use	3.23e+03
256	4	13.8	10	34	W65748	Fibronectin binding p	3.23e+03	329	4	13.8	43	27	W42222	Degenerate 6-mer pept	3.23e+03
257	4	13.8	10	34	W65747	Fibronectin binding p	3.23e+03	330	4	13.8	45	9	R54446	Amb a I.1 residues 57	3.23e+03
258	4	13.8	10	29	W37247	Angiotensin IV activi	3.23e+03	331	4	13.8	46	39	W88513	Amphotropic hypervari	3.23e+03
259	4	13.8	10	22	W21558	Corticotropin releasi	3.23e+03	332	4	13.8	46	39	W88522	Amphotropic hypervari	3.23e+03
260	4	13.8	11	38	W73368	Human PAF-AH protein	3.23e+03	333	4	13.8	47	9	R54445	Amb a I.1 residues 33	3.23e+03
261	4	13.8	11	19	W02438	RAE 40.1-6 comprisng	3.23e+03	334	4	13.8	58	8	R42865	Platelet aggregation	3.23e+03
262	4	13.8	12	39	W89524	Human latent transfer	3.23e+03	335	4	13.8	58	8	R42864	Platelet aggregation	3.23e+03
263	4	13.8	12	25	W32978	HIV gp41 protein epit	3.23e+03	336	4	13.8	58	8	R42858	Platelet aggregation	3.23e+03
264	4	13.8	12	8	R48300	Peptide fragment whic	3.23e+03	337	4	13.8	60	39	W88509	Hypervariable polypro	3.23e+03
265	4	13.8	13	18	R77970	pV1-M residues 1094-1	3.23e+03	338	4	13.8	63	33	W68284	Androctonus venom ana	3.23e+03
266	4	13.8	13	16	R95170	Influenza colonisatio	3.23e+03	339	4	13.8	69	13	R62745	Immunogenic gastrin p	3.23e+03
267	4	13.8	14	39	W89523	Human latent transfer	3.23e+03	340	4	13.8	71	38	W82563	Rice glycoengin clone	3.23e+03
268	4	13.8	15	38	R82540	HIV-1 p24 epitope pep	3.23e+03	341	4	13.8	71	39	W67907	Human secreted protei	3.23e+03
269	4	13.8	15	18	W03417	CDR3-like constrained	3.23e+03	342	4	13.8	74	23	W10031	Protein encoded by cl	3.23e+03
270	4	13.8	16	13	R62730	Invasin domain for us	3.23e+03	343	4	13.8	76	39	W80979	Variable heavy region	3.23e+03
271	4	13.8	17	38	W85992	CD4-variant generated	3.23e+03	344	4	13.8	76	39	W80982	Variable heavy region	3.23e+03
272	4	13.8	17	39	W89522	Human latent transfer	3.23e+03	345	4	13.8	80	32	W60588	Human Hel-N1 RNA reco	3.23e+03
273	4	13.8	17	23	W24037	VL-A13 peptide.	3.23e+03	346	4	13.8	80	32	W60587	Drosophila K3 RNA reco	3.23e+03
274	4	13.8	17	23	W24038	VL-A15 peptide.	3.23e+03	347	4	13.8	85	3	R14458	HIV-1 hxb2 gag 262-33	3.23e+03
275	4	13.8	17	23	W24036	VL-A12 peptide.	3.23e+03	348	4	13.8	85	3	R14008	HIV-1 hxb2 gag 262-33	3.23e+03
276	4	13.8	17	10	R46997	HLA-DR alpha chain po	3.23e+03	349	4	13.8	86	39	W95377	IL-1 alpha propiece p	3.23e+03
277	4	13.8	18	23	W24034	VL-SH peptide.	3.23e+03	350	4	13.8	91	39	W84456	Zif268 three finger p	3.23e+03
278	4	13.8	18	23	W24046	VL-A15 peptide varian	3.23e+03	351	4	13.8	93	31	W28315	Staphylococcus aureus	3.23e+03
279	4	13.8	18	7	R38274	Reo3R-binding peptide	3.23e+03	352	4	13.8	93	31	R82815	Thermostable inorgani	3.23e+03
280	4	13.8	20	39	W87821	Epitope of a Bcl-2 as	3.23e+03	353	4	13.8	95	19	R98152	Heavy chain IGF from	3.23e+03
281	4	13.8	20	29	W36609	Human E-Cadherin vari	3.23e+03	354	4	13.8	95	35	W64321	Mycobacterium tubercu	3.23e+03
282	4	13.8	20	29	W40255	Human wild-type E-Cad	3.23e+03	355	4	13.8	96	39	W80975	Heavy chain IGF from	3.23e+03
283	4	13.8	20	29	W42127	T-cell epitope peptid	3.23e+03	356	4	13.8	97	39	W80971	Heavy chain IGF from	3.23e+03
284	4	13.8	22	39	W89519	Human latent transfer	3.23e+03	357	4	13.8	97	39	W80973	Heavy chain IGF from	3.23e+03
285	4	13.8	23	39	W89518	Human latent transfer	3.23e+03	358	4	13.8	97	39	W80965	Heavy chain IGF from	3.23e+03
286	4	13.8	24	39	W89525	Human latent transfer	3.23e+03	359	4	13.8	98	39	W80963	Heavy chain immunoglo	3.23e+03
287	4	13.8	24	30	W49950	Conantokin peptide de	3.23e+03	360	4	13.8	101	38	W82491	Ehrlichia sp. E46.3 p	3.23e+03
288	4	13.8	25	38	W82521	HIV-1 p24 epitope pep	3.23e+03	361	4	13.8	101	20	R92318	Chemokine-like protei	3.23e+03
289	4	13.8	25	38	W82522	HIV-1 p24 epitope pep	3.23e+03	362	4	13.8	101	7	R34268	Mutated human VH6 gen	3.23e+03
290	4	13.8	25	22	W17909	Peptide PCF0166 from	3.23e+03	363	4	13.8	103	38	W88579	Secreted protein enco	3.23e+03
291	4	13.8	26	35	W75966	Ecdysis triggering ho	3.23e+03	364	4	13.8	104	31	W54040	Human interferon-indu	3.23e+03
292	4	13.8	26	17	R52422	Mouse heavy chain sur	3.23e+03	365	4	13.8	105	31	W28131	Puruvate dehydrogenase	3.23e+03
293	4	13.8	26	17	R52509	Human heavy chain sur	3.23e+03	366	4	13.8	107	39	W73530	Humanised OKT3 variab	3.23e+03
294	4	13.8	26	17	R52427	Mouse heavy chain sur	3.23e+03	367	4	13.8	107	39	W73528	Humanised OKT3 variab	3.23e+03
295	4	13.8	28	28	W47616	Antimicrobial peptide	3.23e+03	368	4	13.8	107	17	R52032	Light chain variable	3.23e+03
296	4	13.8	28	28	W47634	Antimicrobial peptide	3.23e+03	369	4	13.8	109	38	W82982	Human fibronectin III	3.23e+03
297	4	13.8	29	19	W04493	Staphylococcal entero	3.23e+03	370	4	13.8	110	6	R33691	Variant IGE - mutant	3.23e+03
298	4	13.8	30	34	W60660	Sequence-specific DNA	3.23e+03	371	4	13.8	110	6	R32025	Variant IGE - mutant	3.23e+03
299	4	13.8	33	38	W89053	Sequence ID #753 from	3.23e+03	372	4	13.8	111	4	R22574	Heavy chain VH14.1 fr	3.23e+03
300	4	13.8	33	38	W89055	Sequence ID #755 from	3.23e+03	373	4	13.8	112	28	W36112	Human T-cell receptor	3.23e+03
301	4	13.8	33	38	W89052	Polypeptide fragment	3.23e+03	374	4	13.8	113	38	W89348	Mouse 23F2G heavy cha	3.23e+03
302	4	13.8	33	38	W89056	Sequence ID #756 from	3.23e+03	375	4	13.8	113	9	R47932	Heavy chain region of	3.23e+03
303	4	13.8	33	38	W89054	Sequence ID #754 from	3.23e+03	376	4	13.8	113	8	R38722	Vbeta17 variable regi	3.23e+03
304	4	13.8	33	39	W89281	Granulocyte Ehrlich	3.23e+03	377	4	13.8	114	39	W87708	A cysteine rich solub	3.23e+03
305	4	13.8	34	38	W82989	Human fibronectin III	3.23e+03	378	4	13.8	116	16	R80187	MisP1-derived polypep	3.23e+03
306	4	13.8	34	7	R36583	Virus neutralising ep	3.23e+03	379	4	13.8	117	18	R88715	Human antibody heavy	3.23e+03
307	4	13.8	34	31	W37193	Fibronectin peptide f	3.23e+03	380	4	13.8	117	33	W62812	Amino acid sequence o	3.23e+03
308	4	13.8	35	38	W88597	Secreted protein enco	3.23e+03	381	4	13.8	117	15	R85238	Humanised antibody va	3.23e+03

382	4	13.8	118 39	W95355	Human interleukin (IL	3.23e+03	455	4	13.8	179 38	W89820	Protein encoded by c1	3.23e+03
383	4	13.8	118 33	W62811	Amino acid sequence o	3.23e+03	456	4	13.8	180 38	W85617	Murine zcyto7.	3.23e+03
384	4	13.8	119 3	W62795	Amino acid sequence o	3.23e+03	457	4	13.8	181 11	R58805	Fibronectin binding d	3.23e+03
385	4	13.8	119 6	R30143	MAB GAH variable regi	3.23e+03	458	4	13.8	181 18	R91200	DL-D4 fibronectin bin	3.23e+03
386	4	13.8	119 22	W01584	Lead binding WAB 13D1	3.23e+03	459	4	13.8	184 21	W15415	Napin encoded by clon	3.23e+03
387	4	13.8	120 38	W41402	Humanised antibody 80	3.23e+03	460	4	13.8	185 4	P40108	Sequence of human pre	3.23e+03
388	4	13.8	120 39	W84346	Protein N of PRRSV is	3.23e+03	461	4	13.8	185 4	P40155	Sequence of human pre	3.23e+03
389	4	13.8	123 25	W27556	Human AB heavy chain	3.23e+03	462	4	13.8	188 38	W89818	Protein encoded by c1	3.23e+03
390	4	13.8	123 18	W03718	Anti-rhesus D monoclo	3.23e+03	463	4	13.8	189 38	W85096	Thyroid hormone recep	3.23e+03
391	4	13.8	125 29	W52214	Antibody LDL-52-VH ch	3.23e+03	464	4	13.8	190 39	W89897	Antigen 3 from cluste	3.23e+03
392	4	13.8	126 3	R14154	Ig light chain V regi	3.23e+03	465	4	13.8	193 38	W73365	Macaque PAF-AH protei	3.23e+03
393	4	13.8	128 39	W84335	Protein N of PRRSV is	3.23e+03	466	4	13.8	197 39	W73592	Factor xa inhibitor p	3.23e+03
394	4	13.8	128 39	W84336	Protein N of PRRSV is	3.23e+03	467	4	13.8	197 32	W60736	Xylanase A of Schizop	3.23e+03
395	4	13.8	128 36	W68456	PRRSV isolate Toledo	3.23e+03	468	4	13.8	197 1	P82978	Human SP5 protein.	3.23e+03
396	4	13.8	130 32	W82666	Amino acid sequence o	3.23e+03	469	4	13.8	201 15	R85589	Fc(epsilon) CH2'-CH4	3.23e+03
397	4	13.8	135 38	W82266	T. pallidum 17 kDa li	3.23e+03	470	4	13.8	202 21	W12721	PhlF gene product.	3.23e+03
398	4	13.8	138 9	R44828	Flatfish interferon-1	3.23e+03	471	4	13.8	204 24	W25770	Human HD53 U1.	3.23e+03
399	4	13.8	139 39	W95376	IL-1 alpha proplepie p	3.23e+03	472	4	13.8	205 39	W87813	A human Bcl-2-beta pr	3.23e+03
400	4	13.8	141 10	R53339	KM50 cell-derived Ig	3.23e+03	473	4	13.8	206 29	W29479	Human histone H1 isof	3.23e+03
401	4	13.8	141 25	W30718	Haemoglobin beta(-D)	3.23e+03	474	4	13.8	208 39	W73628	Human secreted protei	3.23e+03
402	4	13.8	144 38	W73382	Human haemoglobin bet	3.23e+03	475	4	13.8	209 29	W37449	Tick vasoactive amine	3.23e+03
403	4	13.8	144 37	W73177	Fragment of ganglios	3.23e+03	476	4	13.8	215 3	R13384	TGF-Barrier protein f	3.23e+03
404	4	13.8	145 1	R20303	Val(1) to Met, His(2)	3.23e+03	477	4	13.8	216 39	W87816	Human secreted protei	3.23e+03
405	4	13.8	145 1	R05586	Acetohydroxy acid syn	3.23e+03	478	4	13.8	216 32	W28267	Staphylococcus aureus	3.23e+03
406	4	13.8	145 4	R20305	Val(1) to Met, His(2)	3.23e+03	479	4	13.8	220 39	W81640	Mouse elf-3 protein.	3.23e+03
407	4	13.8	146 38	W73391	Human haemoglobin gam	3.23e+03	480	4	13.8	225 37	W82504	Human EPRG1 protein #	3.23e+03
408	4	13.8	146 38	W73392	Human haemoglobin gam	3.23e+03	481	4	13.8	226 38	W86332	Kidney injury associa	3.23e+03
409	4	13.8	146 38	W73388	Human haemoglobin bet	3.23e+03	482	4	13.8	226 20	W09320	Human chymase clone 7	3.23e+03
410	4	13.8	146 38	W73395	Human haemoglobin bet	3.23e+03	483	4	13.8	227 19	W05530	Mouse TRADD intracell	3.23e+03
411	4	13.8	146 38	W73384	Human haemoglobin bet	3.23e+03	484	4	13.8	229 38	W67566	T. reesei xylanase I	3.23e+03
412	4	13.8	146 38	W73393	Human haemoglobin gam	3.23e+03	485	4	13.8	229 22	W09779	Epstein Barr virus-in	3.23e+03
413	4	13.8	146 38	W73381	Human haemoglobin bet	3.23e+03	486	4	13.8	230 29	W55398	H. pylori ORF hp3p108	3.23e+03
414	4	13.8	146 38	W73383	Human haemoglobin bet	3.23e+03	487	4	13.8	232 3	P60177	Sequence encoded by A	3.23e+03
415	4	13.8	146 38	W73394	Human haemoglobin bet	3.23e+03	488	4	13.8	236 39	W87811	A murine Bcl-2 protei	3.23e+03
416	4	13.8	146 38	W73396	Human haemoglobin bet	3.23e+03	489	4	13.8	236 2	R24191	Bovine RSV strain FS-	3.23e+03
417	4	13.8	146 38	W73387	Human haemoglobin bet	3.23e+03	490	4	13.8	236 13	R71477	Leishmania donovani A	3.23e+03
418	4	13.8	146 4	R21916	Met(75) beta-globin m	3.23e+03	491	4	13.8	239 39	W87812	A human Bcl-2-alpha p	3.23e+03
419	4	13.8	146 25	W30717	Recombinant wild type	3.23e+03	492	4	13.8	239 39	W87810	A human Bcl-2 protein	3.23e+03
420	4	13.8	146 4	R20310	Asp(4) beta-globin mu	3.23e+03	493	4	13.8	239 19	W01018	Apoptosis-blocking pr	3.23e+03
421	4	13.8	146 4	R20309	Glu(4) beta-globin mu	3.23e+03	494	4	13.8	239 8	R42312	Bcl-2 oncogene produc	3.23e+03
422	4	13.8	146 25	W30723	Haemoglobin beta chal	3.23e+03	495	4	13.8	240 7	R40751	Sequence of haemolysi	3.23e+03
423	4	13.8	147 29	W37248	Rat EX clone protein	3.23e+03	496	4	13.8	240 19	W04272	Treponehma hyodysente	3.23e+03
424	4	13.8	149 39	W95375	IL-1 alpha proplepie p	3.23e+03	497	4	13.8	241 22	W20601	H. pylori membrane pr	3.23e+03
425	4	13.8	149 6	R31620	Restrictocin analogue	3.23e+03	498	4	13.8	241 2	R24185	Bovine RSV strain A 5	3.23e+03
426	4	13.8	151 39	W89785	Staphylococcus aureus	3.23e+03	499	4	13.8	244 38	W83019	A murine OCIF-binding	3.23e+03
427	4	13.8	151 17	R85316	Human retinoic acid r	3.23e+03	500	4	13.8	244 34	W89996	Primate chemokine rec	3.23e+03
428	4	13.8	154 21	W00195	Mutant eukaryotic ini	3.23e+03	501	4	13.8	245 38	W88654	Secreted protein enco	3.23e+03
429	4	13.8	154 21	W00197	Mutant eukaryotic ini	3.23e+03	502	4	13.8	246 38	W83020	Osteoclastogenesis in	3.23e+03
430	4	13.8	154 4	R21743	LECGNA3.	3.23e+03	503	4	13.8	246 39	W95194	Human PMM2 protein.	3.23e+03
431	4	13.8	155 14	R79968	Up to the fourth tran	3.23e+03	504	4	13.8	247 3	R14031	Human BDNF.	3.23e+03
432	4	13.8	157 37	W82505	Human EPRG1 protein #	3.23e+03	505	4	13.8	248 39	W83370	Streptococcus pneumon	3.23e+03
433	4	13.8	158 38	W85646	Murine zcyto7 mature	3.23e+03	506	4	13.8	248 28	W73741	DR alpha-DAF chimeric	3.23e+03
434	4	13.8	159 4	R21745	LECGNA8.	3.23e+03	507	4	13.8	248 3	R12468	Luifia cylindrica bloa	3.23e+03
435	4	13.8	160 38	W85627	Murine zcyto7 mature	3.23e+03	508	4	13.8	249 39	W86310	Kidney injury associa	3.23e+03
436	4	13.8	160 38	W85645	Pheromone receptor cl	3.23e+03	509	4	13.8	249 38	W89822	Protein encoded by c1	3.23e+03
437	4	13.8	160 5	R27881	S. pneumoniae ribonuclease fr	3.23e+03	510	4	13.8	249 4	R20930	Envelope proteins fro	3.23e+03
438	4	13.8	163 39	W88487	S. pneumoniae protein	3.23e+03	511	4	13.8	250 24	W25144	Monordin (a ribosome	3.23e+03
439	4	13.8	163 36	W80707	Human kringle 2-3.	3.23e+03	512	4	13.8	252 22	W20502	H. pylori surface mem	3.23e+03
440	4	13.8	168 21	W07562	E. coli colonisation	3.23e+03	513	4	13.8	253 19	P41738	Wasp venom 30 kDa ins	3.23e+03
441	4	13.8	168 27	W38342	Human adult testis se	3.23e+03	514	4	13.8	256 1	P81782	Sequence encoded by o	3.23e+03
442	4	13.8	172 39	W82001	AK wild type beta sub	3.23e+03	515	4	13.8	256 1	P90531	Hypodermin B of ATCC	3.23e+03
443	4	13.8	172 10	R51468	AK wild type beta sub	3.23e+03	516	4	13.8	258 33	W55071	Streptococcus pneumon	3.23e+03
444	4	13.8	172 10	R51467	AK beta subunit r30A.	3.23e+03	517	4	13.8	260 39	W87703	A human serine protea	3.23e+03
445	4	13.8	174 39	W84334	Gp4 protein of PRRSV	3.23e+03	518	4	13.8	261 7	R73298	Plant type I RIP Pore	3.23e+03
446	4	13.8	174 39	W95492	M. leprae RP-factor-1	3.23e+03	519	4	13.8	263 38	W85902	Bovine prion protein	3.23e+03
447	4	13.8	174 11	R58808	Fbp fibronectin bindi	3.23e+03	520	4	13.8	263 7	R37296	Plant type I RIP Mono	3.23e+03
448	4	13.8	175 39	W84333	Thermostable inorgani	3.23e+03	521	4	13.8	263 18	R86716	Bovine prion protein,	3.23e+03
449	4	13.8	176 19	R98153	Staphylococcus aureus	3.23e+03	522	4	13.8	265 38	W89987	Expressed antigen for	3.23e+03
450	4	13.8	177 39	W89778	Gp4 protein of PRRSV	3.23e+03	523	4	13.8	268 39	W95360	Goat interleukin (IL)	3.23e+03
451	4	13.8	178 39	W84332	ORF 4 protein of PRRS	3.23e+03	524	4	13.8	268 39	W95367	Sheep interleukin (IL	3.23e+03
452	4	13.8	178 23	W25960	PRRSV VR 2385 ORF-4 p	3.23e+03	525	4	13.8	268 39	W95358	Bovine interleukin (I	3.23e+03
453	4	13.8	178 17	R94721	ORF 4 protein of PRRS	3.23e+03	526	4	13.8	268 22	W20322	H. pylori secreted or	3.23e+03
454	4	13.8	178 23	W25952	ORF 4 protein of PRRS	3.23e+03	527	4	13.8	268 1	P80756	Sequence encoded by b	3.23e+03



528	4	13.8	269 38	W89860	Antigen 1 from cluste	3.23e+03	601	4	13.8	363 39	W89789	Staphylococcus aureus	3.23e+03
529	4	13.8	270 39	W95370	Pig interleukin (IL)-	3.23e+03	602	4	13.8	363 36	W5788	Staphylococcus aureus	3.23e+03
530	4	13.8	270 39	W95364	Feline interleukin (I	3.23e+03	603	4	13.8	364 39	W53536	IL-1 alpha propleice-E	3.23e+03
531	4	13.8	270 39	W95363	Equine interleukin (I	3.23e+03	604	4	13.8	364 29	W55436	H. pylori ORF hp4p134	3.23e+03
532	4	13.8	271 39	W95357	Human interleukin (IL	3.23e+03	605	4	13.8	367 38	W89051	Polypeptide fragment	3.23e+03
533	4	13.8	272 39	W95329	Human interleukin (IL	3.23e+03	606	4	13.8	367 34	W61492	Human fetuin glycopro	3.23e+03
534	4	13.8	275 18	R95571	Insulin-like growth f	3.23e+03	607	4	13.8	370 37	W74845	Human secreted protei	3.23e+03
535	4	13.8	275 19	W40737	Wasp venom Brhix-1 su	3.23e+03	608	4	13.8	372 14	R76782	Beta-delta opiate rece	3.23e+03
536	4	13.8	276 27	W41560	Pseudomonas fluoresce	3.23e+03	609	4	13.8	374 2	R24147	Beta-lytic protease.	3.23e+03
537	4	13.8	281 29	W44354	Human AGP-1.	3.23e+03	610	4	13.8	375 39	W83067	Pyruvate formate lyas	3.23e+03
538	4	13.8	287 37	W68559	S. peucetius dnrU gen	3.23e+03	611	4	13.8	375 6	R33568	CKS-HCV antigen fusio	3.23e+03
539	4	13.8	288 38	W82482	Mouse bispecific anti	3.23e+03	612	4	13.8	375 19	R50938	Maspin.	3.23e+03
540	4	13.8	290 38	W86337	Kidney injury associa	3.23e+03	613	4	13.8	375 19	W02544	Calcineurin deletion	3.23e+03
541	4	13.8	291 1	R05171	Natural killer cell s	3.23e+03	614	4	13.8	376 28	R73759	Psbg (WbpG) protein i	3.23e+03
542	4	13.8	294 38	W86275	Epichlorohydrin epoxi	3.23e+03	615	4	13.8	378 39	W89898	Antigen 4 from cluste	3.23e+03
543	4	13.8	296 28	W46465	Sensory and motor neu	3.23e+03	616	4	13.8	379 27	R33185	Corn barnacle G-prote	3.23e+03
544	4	13.8	299 38	W88550	Secreted protein enco	3.23e+03	617	4	13.8	382 39	W88491	Human liver clone HP0	3.23e+03
545	4	13.8	300 11	R60803	Yellowjacket phosphol	3.23e+03	618	4	13.8	382 39	W87891	Protein encoded by th	3.23e+03
546	4	13.8	301 39	W87740	Corynebacterium dhyd	3.23e+03	619	4	13.8	382 25	W26551	Chimeric receptor hCT	3.23e+03
547	4	13.8	301 16	R48751	G-protein coupled rat	3.23e+03	620	4	13.8	383 32	W58730	Allergen Amb a ID clo	3.23e+03
548	4	13.8	301 19	W02723	G-protein coupled rat	3.23e+03	621	4	13.8	386 11	R60550	Human developmental t	3.23e+03
549	4	13.8	303 33	W61239	Streptococcus pneumon	3.23e+03	622	4	13.8	388 25	W26592	Human bone morphogene	3.23e+03
550	4	13.8	308 25	W32430	Mycobacterium tubercu	3.23e+03	623	4	13.8	389 37	W80805	Amino acid sequence o	3.23e+03
551	4	13.8	309 39	W73639	Mouse B7-2 antigen.	3.23e+03	624	4	13.8	392 39	W67891	Human secreted protei	3.23e+03
552	4	13.8	309 39	W81948	Tumour rejection anti	3.23e+03	625	4	13.8	392 39	W82714	Soil derived peptide	3.23e+03
553	4	13.8	310 39	W83322	Single chain Apo-2 an	3.23e+03	626	4	13.8	393 19	R99416	Amino acid sequence o	3.23e+03
554	4	13.8	310 39	W83324	Single chain Apo-2 an	3.23e+03	627	4	13.8	393 6	R33579	Minopeptidase precu	3.23e+03
555	4	13.8	312 39	W83323	Single chain Apo-2 an	3.23e+03	628	4	13.8	393 7	R33637	HCV CKS-NS51 recombin	3.23e+03
556	4	13.8	313 34	W42640	Protein sequence that	3.23e+03	629	4	13.8	394 3	R13281	Glucose isomerase mut	3.23e+03
557	4	13.8	315 38	W70805	Amino acid sequence o	3.23e+03	630	4	13.8	394 1	R05282	Amino acid sequence o	3.23e+03
558	4	13.8	315 37	W74867	Human secreted protei	3.23e+03	631	4	13.8	394 3	R13279	Glucose isomerase mut	3.23e+03
559	4	13.8	316 38	W83017	Osteoclastogenesis in	3.23e+03	632	4	13.8	395 28	W34308	Mouse neurotactin.	3.23e+03
560	4	13.8	316 38	W83019	Human osteoprotegerin	3.23e+03	633	4	13.8	395 1	P91956	HIV p41 gene with p12	3.23e+03
561	4	13.8	317 38	W83195	Human osteoprotegerin	3.23e+03	634	4	13.8	396 39	W89680	Human osteogenic prot	3.23e+03
562	4	13.8	317 38	W83018	Human osteoprotegerin	3.23e+03	635	4	13.8	396 29	W44303	Human osteogenic prot	3.23e+03
563	4	13.8	317 33	W69557	NR-kB receptor activa	3.23e+03	636	4	13.8	396 23	W18574	Aggrecanase artificia	3.23e+03
564	4	13.8	318 39	W82313	Human A3 adenosine re	3.23e+03	637	4	13.8	396 7	R33631	HCV CKS-Core pHCV34 a	3.23e+03
565	4	13.8	319 39	W82724	Adenovirus T5S7sigDEL	3.23e+03	638	4	13.8	398 36	W71071	Multiple sclerosis as	3.23e+03
566	4	13.8	321 39	W83375	Streptococcus pneumon	3.23e+03	639	4	13.8	399 39	W82711	Soil derived peptide	3.23e+03
567	4	13.8	322 2	P81094	Sequence of rhinoviru	3.23e+03	640	4	13.8	399 39	W95505	Mortierella alpina de	3.23e+03
568	4	13.8	324 35	W11525	Helicobacter polysept	3.23e+03	641	4	13.8	400 36	W72648	Canine herpes virus p	3.23e+03
569	4	13.8	327 23	W08483	Sequence of ovine IL-	3.23e+03	642	4	13.8	402 1	R05277	New serine hydroxymet	3.23e+03
570	4	13.8	328 30	W53247	Mouse haematopoietic	3.23e+03	643	4	13.8	404 30	W53456	Protein 41spf-1 for e	3.23e+03
571	4	13.8	328 29	W44324	Heterodimeric human I	3.23e+03	644	4	13.8	404 31	W54077	LR-2 protein #1.	3.23e+03
572	4	13.8	328 29	W55486	H. pylori ORF hp2p102	3.23e+03	645	4	13.8	407 17	R92312	Alpha-amylase precurs	3.23e+03
573	4	13.8	329 39	W86356	Partial human DNAX to	3.23e+03	646	4	13.8	411 39	W83358	Mycobacterium tubercu	3.23e+03
574	4	13.8	329 32	W58396	Homo sapiens HLIM-2 p	3.23e+03	647	4	13.8	411 13	R78740	Fusion of BMP-2 prope	3.23e+03
575	4	13.8	334 38	W88633	Secreted protein enco	3.23e+03	648	4	13.8	417 22	W18669	Polliovirus receptor (	3.23e+03
576	4	13.8	334 36	W72171	HSV-2 strain S85 Cont	3.23e+03	649	4	13.8	421 10	R51466	AK wild type alpha su	3.23e+03
577	4	13.8	337 34	W69997	Primate chemokine rec	3.23e+03	650	4	13.8	421 12	R63571	Feedback inhibition r	3.23e+03
578	4	13.8	338 38	W82407	E. chrysanthemi HrpN-	3.23e+03	651	4	13.8	421 10	R51465	AK alpha subunit T279	3.23e+03
579	4	13.8	338 39	W87638	A hypersensitive resp	3.23e+03	652	4	13.8	421 15	R86307	Corynebacterium gluta	3.23e+03
580	4	13.8	338 5	R27667	Soya beta-1,3-glucana	3.23e+03	653	4	13.8	421 1	P82725	Ty-HIV fusion protein	3.23e+03
581	4	13.8	338 6	R33450	NSI-27 fusion protei	3.23e+03	654	4	13.8	423 38	W74815	Human secreted protei	3.23e+03
582	4	13.8	338 36	W72151	HSV-2 strain S85 Cont	3.23e+03	655	4	13.8	429 39	W73631	Human secreted protei	3.23e+03
583	4	13.8	339 22	W02906	H. pylori secreted or	3.23e+03	656	4	13.8	432 38	W82246	Bacteriophage fd-tet	3.23e+03
584	4	13.8	340 19	W02549	Calcineurin deletion	3.23e+03	657	4	13.8	432 22	W20733	H. pylori cell envelo	3.23e+03
585	4	13.8	341 2	R08522	Tapetum-specific prom	3.23e+03	658	4	13.8	436 21	W03662	Human 70K u1 snRNP pr	3.23e+03
586	4	13.8	342 34	W68141	Human herpesvirus 8 G	3.23e+03	659	4	13.8	437 31	W03622	P-selectin ligand pro	3.23e+03
587	4	13.8	347 19	W02698	G-protein coupled mou	3.23e+03	660	4	13.8	439 38	W82645	Enrichia sp. extende	3.23e+03
588	4	13.8	348 37	W81607	Ankyrin protein fragm	3.23e+03	661	4	13.8	439 1	P93333	Sequence of C-myc pro	3.23e+03
589	4	13.8	352 36	W71070	Multiple sclerosis as	3.23e+03	662	4	13.8	441 38	W73359	Human PAF-AH protein	3.23e+03
590	4	13.8	352 21	W07576	Murine kringle 1-4.	3.23e+03	663	4	13.8	441 14	R71923	Human acetyl hydrolas	3.23e+03
591	4	13.8	353 30	W33627	Yeast transcriptional	3.23e+03	664	4	13.8	441 14	R71924	Human acetyl hydrolas	3.23e+03
592	4	13.8	353 17	R88413	High-affinity melaton	3.23e+03	665	4	13.8	443 39	W83372	Streptococcus pneumon	3.23e+03
593	4	13.8	354 39	W86311	Kidney injury associa	3.23e+03	666	4	13.8	446 39	R83374	Streptococcus pneumon	3.23e+03
594	4	13.8	354 21	W14463	Yeast glycoprotein (Y	3.23e+03	667	4	13.8	448 1	P90341	hap (hepatoma) protei	3.23e+03
595	4	13.8	355 21	W11475	Marek's disease virus	3.23e+03	668	4	13.8	448 13	R87664	Human RAR-beta.	3.23e+03
596	4	13.8	360 38	W70804	Amino acid sequence o	3.23e+03	669	4	13.8	456 38	R82657	Ehrlichia sp. extende	3.23e+03
597	4	13.8	360 17	R94375	Laminin A chain (aa23	3.23e+03	670	4	13.8	456 36	W69436	Human phospholipase A	3.23e+03
598	4	13.8	361 18	R91621	Human immunodeficienc	3.23e+03	671	4	13.8	456 13	R74996	E. maxima Em70-1 anti	3.23e+03
599	4	13.8	362 1	R04211	Deduced amino acid se	3.23e+03	672	4	13.8	457 39	W73633	Human secreted protei	3.23e+03
600	4	13.8	362 8	R41536	Preprokiller 2 toxin	3.23e+03	673	4	13.8	460 2	P93629	Sequence of human pro	3.23e+03

674	4	13.8	464 39	W84184	Glial cell line-deriv	3.23e+03	747	4	13.8	548 36	W72252	HSV-2 strain SB5 Cont	3.23e+03
675	4	13.8	464 39	W84179	A GNFR-alpha-related	3.23e+03	748	4	13.8	549 39	W67799	Thermococcus sp. KS-8	3.23e+03
676	4	13.8	464 8	R42906	Human antithrombin II	3.23e+03	749	4	13.8	550 14	R66627	Firefly luciferase (A	3.23e+03
677	4	13.8	465 22	W12666	Beta-amyloid precursor	3.23e+03	750	4	13.8	551 39	W88488	Rat organic anion tra	3.23e+03
678	4	13.8	466 38	W82398	Mouse G3BP protein.	3.23e+03	751	4	13.8	551 39	W73480	Grapevine leafroll vi	3.23e+03
679	4	13.8	466 19	R95688	Human GAP-SH3 domain	3.23e+03	752	4	13.8	551 30	W55884	Human CD33-like prote	3.23e+03
680	4	13.8	467 39	W82595	Mouse Smad2 protein.	3.23e+03	753	4	13.8	551 25	W32937	Modified firefly luci	3.23e+03
681	4	13.8	467 38	W82614	Murine Smad2 S464A/S4	3.23e+03	754	4	13.8	552 32	W48873	Hyphozyma sp. strain	3.23e+03
682	4	13.8	467 38	W82613	Murine Smad2 S465A/S4	3.23e+03	755	4	13.8	555 4	R22925	Truncated hamster HMG	3.23e+03
683	4	13.8	467 38	W82609	Murine Smad2 S465A mu	3.23e+03	756	4	13.8	557 29	W43448	Tobacco laccase clone	3.23e+03
684	4	13.8	467 38	W82607	Human Smad2 protein.	3.23e+03	757	4	13.8	558 38	W73379	CSB2 protein sequence	3.23e+03
685	4	13.8	467 38	W82608	Murine Smad2 S464A mu	3.23e+03	758	4	13.8	559 39	W89803	Staphylococcus aureus	3.23e+03
686	4	13.8	467 38	W82611	Murine Smad2 S464A/S4	3.23e+03	759	4	13.8	560 38	W79954	Truncated adenovirus	3.23e+03
687	4	13.8	467 38	W82612	Murine Smad2 S464A/S4	3.23e+03	760	4	13.8	560 17	R98763	Glial growth factor I	3.23e+03
688	4	13.8	467 38	W82610	Murine Smad2 S467A mu	3.23e+03	761	4	13.8	561 19	R91932	Tripartite growth factor	3.23e+03
689	4	13.8	468 30	W55675	H. pylori ORF hp06122	3.23e+03	762	4	13.8	562 39	W85602	Phosphatidylinositol-	3.23e+03
690	4	13.8	473 3	R14195	Rat cannabinoid recep	3.23e+03	763	4	13.8	563 39	W88489	Hexasaminidase enzyme	3.23e+03
691	4	13.8	474 31	W58862	T. halophilus xylose	3.23e+03	764	4	13.8	568 38	W79953	Human organic anion t	3.23e+03
692	4	13.8	475 3	R14625	S. typhimurium HtrA p	3.23e+03	765	4	13.8	568 34	W62989	Truncated adenovirus	3.23e+03
693	4	13.8	477 2	R05122	Bat-PA(H).	3.23e+03	766	4	13.8	569 15	R84923	HN protein of Newcastle	3.23e+03
694	4	13.8	478 9	R47586	Human bone formation-	3.23e+03	767	4	13.8	572 13	R78520	Spinach 2-oxoglutarat	3.23e+03
695	4	13.8	479 29	W55629	H. pylori ORF 02ae310	3.23e+03	768	4	13.8	573 38	W79955	Partial ALK protein.	3.23e+03
696	4	13.8	479 32	P92606	Human muscarinic acet	3.23e+03	769	4	13.8	573 38	W79952	Truncated adenovirus	3.23e+03
697	4	13.8	479 9	R46235	A. niger pH 2.5 acid	3.23e+03	770	4	13.8	574 38	W81559	Respiratory syncytial	3.23e+03
698	4	13.8	480 36	W72251	HSV-2 strain SB5 Cont	3.23e+03	771	4	13.8	574 29	W47604	HRSV glycoprotein F.	3.23e+03
699	4	13.8	484 39	W82974	Granulocytic Ehrlich	3.23e+03	772	4	13.8	576 39	W88523	EAT-4 protein amino a	3.23e+03
700	4	13.8	484 23	W22169	S. thermophilus exopol	3.23e+03	773	4	13.8	581 38	W79949	Mutant adenovirus typ	3.23e+03
701	4	13.8	487 33	W55047	Alkaline phosphatase	3.23e+03	774	4	13.8	581 38	W79950	Mutant adenovirus typ	3.23e+03
702	4	13.8	490 20	W10207	Human p33 binding pro	3.23e+03	775	4	13.8	581 38	W79950	Adenovirus type fibre	3.23e+03
703	4	13.8	493 4	P40065	Sequence of human imm	3.23e+03	776	4	13.8	581 39	W82732	Adenovirus wild-type	3.23e+03
704	4	13.8	493 17	R81467	Human derived cytochr	3.23e+03	777	4	13.8	581 38	W5419	Mutant adenovirus typ	3.23e+03
705	4	13.8	494 38	W73350	Rabbit alaphad protein	3.23e+03	778	4	13.8	581 38	W79951	Mutant adenovirus typ	3.23e+03
706	4	13.8	494 29	W47312	Zucchini ACC synthase	3.23e+03	779	4	13.8	584 2	P94202	Sequence of alpha-glu	3.23e+03
707	4	13.8	496 37	W75589	Simian immunodeficien	3.23e+03	780	4	13.8	588 36	W71881	Anti-human Fas human	3.23e+03
708	4	13.8	496 34	W76984	Mouse Egr-1 protein.	3.23e+03	781	4	13.8	590 39	W84265	Protein encoded by cp	3.23e+03
709	4	13.8	496 34	W70005	Atrazine chlorohydrol	3.23e+03	782	4	13.8	591 33	W69950	RANK polypeptide part	3.23e+03
710	4	13.8	496 34	W70006	Atrazine chlorohydrol	3.23e+03	783	4	13.8	592 38	W70797	Human interleukin-6R-	3.23e+03
711	4	13.8	499 37	W76295	Polyporus pinus (I)	3.23e+03	784	4	13.8	594 6	R33566	CKS-HCV antigen fusio	3.23e+03
712	4	13.8	499 37	W76297	Polyporus pinus (I)	3.23e+03	785	4	13.8	595 38	W87487	S. cerevisiae TH1 po	3.23e+03
713	4	13.8	499 37	W76298	Polyporus pinus (I)	3.23e+03	786	4	13.8	598 2	R12319	HIV-2/HIV-1/HTLV-I ch	3.23e+03
714	4	13.8	500 36	W70562	HIV-1 GAG/PRO polyep	3.23e+03	787	4	13.8	599 39	W73479	Grapevine leafroll vi	3.23e+03
715	4	13.8	500 33	W63744	HIV-1 NL-43 gag prote	3.23e+03	788	4	13.8	605 1	R06239	Novel mutant tissue p	3.23e+03
716	4	13.8	500 33	W63745	HIV-1 NL-43 gag prote	3.23e+03	789	4	13.8	613 31	W37799	Amino acid sequence o	3.23e+03
717	4	13.8	502 18	R96087	Yeast calcineurin sub	3.23e+03	790	4	13.8	613 6	R33567	CKS-HCV antigen fusio	3.23e+03
718	4	13.8	503 31	W53232	Cryptosporidium parvu	3.23e+03	791	4	13.8	616 31	W55980	Aspergillus awamori g	3.23e+03
719	4	13.8	504 4	R21835	Sequence encoded by h	3.23e+03	792	4	13.8	619 39	W89271	Granulocytic Ehrlich	3.23e+03
720	4	13.8	506 39	W83126	Prf1K48 Lysine speci	3.23e+03	793	4	13.8	620 22	W14993	Human c-fos induced g	3.23e+03
721	4	13.8	509 39	W95196	Bcl-2 interaction pro	3.23e+03	794	4	13.8	621 31	W55978	Aspergillus awamori g	3.23e+03
722	4	13.8	511 36	W72141	HSV-2 strain SB5 Cont	3.23e+03	795	4	13.8	622 38	W67474	Murine ecotropic retr	3.23e+03
723	4	13.8	513 15	R79945	Helicobacter pylori a	3.23e+03	796	4	13.8	625 38	W83200	Murine osteoclast dif	3.23e+03
724	4	13.8	517 22	W17975	Coprinus cinereus loc	3.23e+03	797	4	13.8	626 39	W73533	MEKK3 protein.	3.23e+03
725	4	13.8	520 14	R72610	Human calcium channel	3.23e+03	798	4	13.8	630 38	W85018	Jnk1-green flourescen	3.23e+03
726	4	13.8	521 24	W26359	Mouse activin recepto	3.23e+03	799	4	13.8	630 3	R12230	TrpE/androgen recepto	3.23e+03
727	4	13.8	524 39	W95521	Protease activated pr	3.23e+03	800	4	13.8	633 38	W85009	Jnk1-green flourescen	3.23e+03
728	4	13.8	525 13	R78521	NPM/ALK fusion protei	3.23e+03	801	4	13.8	638 1	R05589	Acetohydroxy acid syn	3.23e+03
729	4	13.8	527 38	W81560	Truncated RSV F prote	3.23e+03	802	4	13.8	646 10	R35462	Peptides encoded by t	3.23e+03
730	4	13.8	529 38	W85055	Human protein designa	3.23e+03	803	4	13.8	647 38	W73376	Human HspDV78 protein	3.23e+03
731	4	13.8	529 38	W80498	A protein designated	3.23e+03	804	4	13.8	647 4	R20813	B lymphocyte-specific	3.23e+03
732	4	13.8	529 3	R13503	HSF.	3.23e+03	805	4	13.8	648 23	R24473	Herbicide-sensitive a	3.23e+03
733	4	13.8	530 33	W55046	Human placental alkali	3.23e+03	806	4	13.8	648 4	R22563	Mouse mutant c-rafi-1	3.23e+03
734	4	13.8	531 22	W20739	H. pylori cytoplasmic	3.23e+03	807	4	13.8	648 22	W17045	Mouse c-rafi 1 protein	3.23e+03
735	4	13.8	532 30	W50033	Human immunity relate	3.23e+03	808	4	13.8	648 21	W12777	Bti truncated delta-e	3.23e+03
736	4	13.8	532 32	W58551	Human neutral amino a	3.23e+03	809	4	13.8	657 39	W85452	Pig endogenous retrov	3.23e+03
737	4	13.8	533 2	R12276	Recombinant aspartate	3.23e+03	810	4	13.8	659 29	W37744	CD2 associated intrac	3.23e+03
738	4	13.8	534 35	W71223	63 KDa Cam-PDE encode	3.23e+03	811	4	13.8	660 39	W85453	Pig endogenous retrov	3.23e+03
739	4	13.8	536 37	W79272	Phenolic acid esteras	3.23e+03	812	4	13.8	664 35	W71603	Human neurturin recep	3.23e+03
740	4	13.8	537 39	W87796	Tripeptidyl aminopept	3.23e+03	813	4	13.8	664 3	P60229	Alcohol-oxidase.	3.23e+03
741	4	13.8	538 38	W89628	Secreted protein enco	3.23e+03	814	4	13.8	667 5	R26912	ALS C3 mutant of toba	3.23e+03
742	4	13.8	540 39	W89794	Staphylococcus aureus	3.23e+03	815	4	13.8	668 5	R24801	pro197/Ser653 deleted	3.23e+03
743	4	13.8	540 39	W89777	Staphylococcus aureus	3.23e+03	816	4	13.8	668 28	W47037	Soluble truncated VEG	3.23e+03
744	4	13.8	540 35	W77145	Human interleukin-18	3.23e+03	817	4	13.8	669 32	W61952	Murine mortalin amino	3.23e+03
745	4	13.8	541 22	W05767	Presenilin homologue.	3.23e+03	818	4	13.8	673 39	W89587	SpHINGOMONAS capsulat	3.23e+03
746	4	13.8	542 39	W73634	Human secreted protei	3.23e+03	819	4	13.8	675 35	W69354	P. damsela bst protei	3.23e+03

820	4	13.8	676	28	W46464	Human protein S.	3.23e+03	893	4	13.8	880	38	W73332	HIV envelope protein.	3.23e+03
821	4	13.8	677	20	W06083	Human TARA-binding pr	3.23e+03	894	4	13.8	887	18	R95004	CagC antigenic polyope	3.23e+03
822	4	13.8	679	33	W55041	Homo sapiens LH receptor	3.23e+03	895	4	13.8	891	38	W85583	Human N-methyl-D-aspa	3.23e+03
823	4	13.8	687	28	W47036	Soluble VEGF receptor	3.23e+03	896	4	13.8	897	39	W73474	Human retinoblastoma	3.23e+03
824	4	13.8	688	17	R65331	Paralichthys olivaceu	3.23e+03	897	4	13.8	896	23	W69371	Modified retinoblasto	3.23e+03
825	4	13.8	688	17	R39519	Transglutaminase (fis	3.23e+03	898	4	13.8	907	35	W16313	Yeast alpha factor-G-	3.23e+03
826	4	13.8	689	6	R30309	N-terminal of LH rece	3.23e+03	899	4	13.8	912	28	W08959	Amino acid sequence o	3.23e+03
827	4	13.8	693	36	W80722	S. pneumoniae elongat	3.23e+03	900	4	13.8	918	2	R10545	Recombinant human gpl	3.23e+03
828	4	13.8	698	39	W83493	4D5 Fab molecule expr	3.23e+03	901	4	13.8	918	9	R46233	Human soluble glycopr	3.23e+03
829	4	13.8	700	6	R30693	N-terminal of LH rece	3.23e+03	902	4	13.8	920	9	R49043	NMDA receptor channel	3.23e+03
830	4	13.8	704	20	W06085	Human TARA-binding pr	3.23e+03	903	4	13.8	928	1	R06289	Predicted retinoblast	3.23e+03
831	4	13.8	711	37	W73025	Helicobacter pylori 7	3.23e+03	904	4	13.8	938	38	R85572	Human N-methyl-D-aspa	3.23e+03
832	4	13.8	711	37	R42845	Mutant cyclonulcodext	3.23e+03	905	4	13.8	944	39	W83318	Mouse SMAD interactin	3.23e+03
833	4	13.8	714	23	W22377	S. pneumoniae FucI/HS	3.23e+03	906	4	13.8	945	39	W73624	Human secreted protei	3.23e+03
834	4	13.8	716	33	W69846	Amino acid sequence o	3.23e+03	907	4	13.8	947	39	W81168	Transcriptional regul	3.23e+03
835	4	13.8	716	33	W44896	PreproHCPB-linker-hum	3.23e+03	908	4	13.8	948	38	W81168	Rat receptor tyrosine	3.23e+03
836	4	13.8	718	38	W85024	EtK2-green fluorescen	3.23e+03	909	4	13.8	950	39	W83319	Mouse SMAD interactin	3.23e+03
837	4	13.8	719	38	W85012	Snad2-green fluoresce	3.23e+03	910	4	13.8	951	38	W70798	Human gpl30-C-gamma-1	3.23e+03
838	4	13.8	719	11	R60193	Modified protective a	3.23e+03	911	4	13.8	959	38	W85582	Human N-methyl-D-aspa	3.23e+03
839	4	13.8	720	16	R92227	Thermostable enzyme.	3.23e+03	912	4	13.8	968	28	W41309	CF-5 pathogen resista	3.23e+03
840	4	13.8	722	36	W73107	B.t. toxin protein 21	3.23e+03	913	4	13.8	969	24	W25170	Human insulinoma-asso	3.23e+03
841	4	13.8	731	33	W44896	Polyproline-contains	3.23e+03	914	4	13.8	979	39	W83068	TuTC protein of tolu	3.23e+03
842	4	13.8	734	32	W46308	AAV4 VPI capsid prote	3.23e+03	915	4	13.8	988	13	R75710	Eph-related PTK Cekl0	3.23e+03
843	4	13.8	735	11	R60179	Protective antigen of	3.23e+03	916	4	13.8	1005	20	W11238	Beta-galactosidase.	3.23e+03
844	4	13.8	735	11	R54867	Hepatitis C Virus cor	3.23e+03	917	4	13.8	1012	39	W95395	IBDV VP2-VP3-VP4 poly	3.23e+03
845	4	13.8	746	26	W26673	Staphylococcus aureus	3.23e+03	918	4	13.8	1015	27	W35345	Human protein tyrosin	3.23e+03
846	4	13.8	746	33	W44897	Polyproline-contains	3.23e+03	919	4	13.8	1027	22	W20217	H. pylori cytoplasmic	3.23e+03
847	4	13.8	748	39	W89272	Granulocytic Ehrlich	3.23e+03	920	4	13.8	1043	30	W36509	Murine RENT1 protein.	3.23e+03
848	4	13.8	750	39	R82728	Adenovirus PNPi4. SS	3.23e+03	921	4	13.8	1048	4	R20637	Human Cytomegalovirus	3.23e+03
849	4	13.8	753	39	W83927	Human T85 protein.	3.23e+03	922	4	13.8	1058	19	W00362	Human ubiquitin-activ	3.23e+03
850	4	13.8	756	2	R08405	AmF105L encoded by F1	3.23e+03	923	4	13.8	1092	39	W88257	Gracilariaopsis lemane	3.23e+03
851	4	13.8	758	31	W46270	Moraxella catarrhalis	3.23e+03	924	4	13.8	1101	39	W90089	Human G-protein regul	3.23e+03
852	4	13.8	763	32	W60591	Human hepatocyte nucl	3.23e+03	925	4	13.8	1102	39	W90083	Porcine G-protein reg	3.23e+03
853	4	13.8	763	30	W41734	Human TRAF-2 kinase.	3.23e+03	926	4	13.8	1114	36	W72205	HSV-2 strain S85 Cont	3.23e+03
854	4	13.8	772	33	W44898	Polyproline-contains	3.23e+03	927	4	13.8	1120	39	W81642	Mouse elf protein.	3.23e+03
855	4	13.8	780	28	W47039	Soluble truncated VEG	3.23e+03	928	4	13.8	1130	3	R13436	Merosin M polypeptide	3.23e+03
856	4	13.8	781	6	R33374	HCV CKS-33-BCD protei	3.23e+03	929	4	13.8	1144	15	R88122	Tobacco mosaic virus	3.23e+03
857	4	13.8	781	7	R33632	HCV CKS-33-BCD fusion	3.23e+03	930	4	13.8	1155	38	W73346	Mouse alphad protein	3.23e+03
858	4	13.8	784	39	W87760	Soybean lysine ketogl	3.23e+03	931	4	13.8	1155	5	R28047	IRS-1 protein.	3.23e+03
859	4	13.8	786	39	R86349	Human DNAX toll-like	3.23e+03	932	4	13.8	1161	38	W73347	Mouse alphaad protein	3.23e+03
860	4	13.8	788	26	W26672	Staphylococcus aureus	3.23e+03	933	4	13.8	1161	32	W65103	Mouse beta-integrin a	3.23e+03
861	4	13.8	798	30	W36334	Yeast transcripional	3.23e+03	934	4	13.8	1184	6	R32882	Cardiac adenyllyl cycl	3.23e+03
862	4	13.8	802	37	W81746	M. tuberculosis fusio	3.23e+03	935	4	13.8	1220	36	W73106	B.t. toxin protein 86	3.23e+03
863	4	13.8	810	39	W73471	Mouse interleukin-4 r	3.23e+03	936	4	13.8	1228	17	R77673	S-layer protein encod	3.23e+03
864	4	13.8	812	21	W07585	Murine plasminogen se	3.23e+03	937	4	13.8	1265	38	W68594	Human methionine synt	3.23e+03
865	4	13.8	812	3	R12948	Plasminogen miteins T	3.23e+03	938	4	13.8	1267	1	R04232	Rickettsia rickettsii	3.23e+03
866	4	13.8	813	3	R12936	Plasminogen miteins X5	3.23e+03	939	4	13.8	1274	39	W89253	Human ALP	3.23e+03
867	4	13.8	813	3	R12935	Plasminogen miteins X3	3.23e+03	940	4	13.8	1285	37	W72972	Drosophila melanogast	3.23e+03
868	4	13.8	814	3	R12947	Plasminogen miteins T2	3.23e+03	941	4	13.8	1289	6	R28810	BT toxin 17b.	3.23e+03
869	4	13.8	814	8	R43378	Peptide which modulat	3.23e+03	942	4	13.8	1296	17	R95010	C. botulinum type A n	3.23e+03
870	4	13.8	822	38	W73486	Cellulose phosphoryl	3.23e+03	943	4	13.8	1301	25	W29115	FMR2P protein associa	3.23e+03
871	4	13.8	830	17	R81627	Human trkC receptor p	3.23e+03	944	4	13.8	1304	37	W59994	Human neural cell adh	3.23e+03
872	4	13.8	832	35	W69368	Modified retinoblasto	3.23e+03	945	4	13.8	1336	38	W87511	Human N-methyl-D-aspa	3.23e+03
873	4	13.8	832	20	W09256	Bacillus alkaline alp	3.23e+03	946	4	13.8	1356	32	W59275	Human KDR protein.	3.23e+03
874	4	13.8	842	12	R63570	Feedback inhibition r	3.23e+03	947	4	13.8	1367	19	R97420	Murine foetal liver k	3.23e+03
875	4	13.8	849	38	W82397	Human UBP protein #3.	3.23e+03	948	4	13.8	1404	5	R26049	MSF precursor.	3.23e+03
876	4	13.8	850	31	W57445	A. thermophilum thrm	3.23e+03	949	4	13.8	1422	16	R82069	Hepatitis GB virus (H	3.23e+03
877	4	13.8	851	32	W43079	HIV-1 gp120 protein f	3.23e+03	950	4	13.8	1423	36	W71367	Death associated prot	3.23e+03
878	4	13.8	853	32	W43077	HIV-1 gp120 protein f	3.23e+03	951	4	13.8	1423	14	R74205	Human death associate	3.23e+03
879	4	13.8	853	32	W43066	HIV-1 gp120 protein f	3.23e+03	952	4	13.8	1424	1	P91169	Sequence of 740 Arg-1	3.23e+03
880	4	13.8	854	32	W43068	HIV-1 gp120 protein f	3.23e+03	953	4	13.8	1435	13	P80232	P. falciparum SABB.	3.23e+03
881	4	13.8	855	32	W43069	HIV-1 gp120 protein f	3.23e+03	954	4	13.8	1458	1	P80804	Sequence of HIV-1 BRU	3.23e+03
882	4	13.8	856	3	P61514	Sequence of envelope	3.23e+03	955	4	13.8	1481	39	W83299	Human rhds528 protein	3.23e+03
883	4	13.8	856	10	R51248	FIV petalum envelope	3.23e+03	956	4	13.8	1481	39	W33363	Human multidrug resis	3.23e+03
884	4	13.8	856	10	R51253	FIV DUTCH113 envelope	3.23e+03	957	4	13.8	1481	39	W33363	High molecular weight	3.23e+03
885	4	13.8	858	32	W43067	HIV-1 gp120 protein f	3.23e+03	958	4	13.8	1492	27	W34623	Human C3 protein muta	3.23e+03
886	4	13.8	859	38	W07796	Human gpl30-Fc-His6 a	3.23e+03	959	4	13.8	1604	13	R70105	TNF-R-EBA 175 fusion	3.23e+03
887	4	13.8	869	35	W69374	Modified retinoblasto	3.23e+03	960	4	13.8	1657	24	W18822	Human IQGAP1.	3.23e+03
888	4	13.8	870	38	W85585	Human N-methyl-D-aspa	3.23e+03	961	4	13.8	1663	27	W34619	Human C3 protein muta	3.23e+03
889	4	13.8	878	39	W95393	IBDV VPI protein.	3.23e+03	962	4	13.8	1745	36	W76776	D. immitis ankyrin pD	3.23e+03
890	4	13.8	878	39	W95397	IBDV VPI protein.	3.23e+03	963	4	13.8	1745	37	W70608	Full length ankyrin p	3.23e+03
891	4	13.8	878	32	W54373	IBDV RNA-dependent RN	3.23e+03	964	4	13.8	1757	39	W84351	Murine ubiquitin-prot	3.23e+03
892	4	13.8	879	4	R24126	SIVmac239 env gene pr	3.23e+03	965	4	13.8	1891	26	W22610	Platenolide synthase	3.23e+03

966 4 13.8 1938 39 W73553 Full length Pullulana 3.23e+03  
967 4 13.8 2039 32 W56322 Haemophilus paragalli 3.23e+03  
968 4 13.8 2098 16 R65863 Factor-VIII 3.23e+03  
969 4 13.8 2154 39 W81639 Mouse elf-1 protein. 3.23e+03  
970 4 13.8 2161 33 W61337 Human calcium channel 3.23e+03  
971 4 13.8 2124 24 W26357 Human LDL receptor an 3.23e+03  
972 4 13.8 2237 14 R71006 Human neuronal calciu 3.23e+03  
973 4 13.8 2332 2 W71726 Facor VIII:c Variant 3.23e+03  
974 4 13.8 2332 30 W53483 Human factor VIII 3.23e+03  
975 4 13.8 2345 24 W11446 Active Factor VIII:C 3.23e+03  
976 4 13.8 2346 24 W11360 Active Factor VIII:C 3.23e+03  
977 4 13.8 2348 24 W11386 Active Factor VIII:C 3.23e+03  
978 4 13.8 2350 24 W11380 Active Factor VIII:C 3.23e+03  
979 4 13.8 2350 24 W11470 Active Factor VIII:C 3.23e+03  
980 4 13.8 2351 24 W11419 Active Factor VIII:C 3.23e+03  
981 4 13.8 2351 31 W45245 Human factor III prot 3.23e+03  
982 4 13.8 2351 24 W11416 Active Factor VIII:C 3.23e+03  
983 4 13.8 2351 24 W10591 Factor VIII:C (Phe652 3.23e+03  
984 4 13.8 2351 24 W11347 Active Factor VIII:C 3.23e+03  
985 4 13.8 2352 24 W11418 Active Factor VIII:C 3.23e+03  
986 4 13.8 2352 24 W11397 Active Factor VIII:C 3.23e+03  
987 4 13.8 2352 24 W11417 Active Factor VIII:C 3.23e+03  
988 4 13.8 2352 24 W11451 Active Factor VIII:C 3.23e+03  
989 4 13.8 2352 24 W11463 Active Factor VIII:C 3.23e+03  
990 4 13.8 2386 34 W63171 Amino acid sequence o 3.23e+03  
991 4 13.8 2441 30 W40058 Cellular transcriptio 3.23e+03  
992 4 13.8 2476 39 W67238 Pig p105 zona pelluci 3.23e+03  
993 4 13.8 2516 19 W01875 Neuronal invertebrate 3.23e+03  
994 4 13.8 2549 32 W56027 FRAP (torl) protein. 3.23e+03  
995 4 13.8 2594 21 W14748 IgG-Fc binding protei 3.23e+03  
996 4 13.8 2639 39 W73476 Grapevine leafroll vi 3.23e+03  
997 4 13.8 3672 27 W31950 Human bg protein asso 3.23e+03  
998 4 13.8 3910 7 R38470 ALL-1 protein. 3.23e+03  
999 4 13.8 4655 29 W43313 Human kidney calcium 3.23e+03  
1000 4 13.8 5405 21 W14749 IgG-Fc binding protei 3.23e+03

## ALIGNMENTS

RESULT 1  
ID W53570 standard; Protein; 784 AA.  
AC W53570;  
DT 06-JUL-1998 (first entry)  
DE Cucumber raffinose synthase.  
KW Cucumber; raffinose synthase; sucrose; galactinol.  
OS Cucumis sativus.  
PN J10084973-A.  
PD 07-APR-1998.  
PF 28-APR-1997; 111124.  
PR 26-JUL-1996; JP-198079.  
PR 26-APR-1996; JP-107682.  
PA (AJIN ) AJINOMOTO KK.  
DR WPI; 98-264858/24.  
PT N-PSDB; V22250.  
PT Raffinose synthase gene - useful for preparation of raffinose in transformed plant  
PS Claim 3; Pages 17-20; 26pp; Japanese.  
CC The present sequence is cucumber raffinose synthase, which forms raffinose from sucrose and galactinol, has an optimum pH of 6 to 8 and working temperature of 35 to 40 degrees C, has a molecular weight of 75 to 95 kDa by gel filtration or 90 to 100 kDa by PAGE and SDS-PAGE under reductive conditions and is inhibited by iodoacetamide, N-ethylmaleimide and myoinositol.  
SQ Sequence 784 AA;

Query Match 20.7%; Score 6; DB 30; Length 784;  
Best Local Similarity 100.0%; Pred. No. 4.45e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 625 vtscn 630  
|||||  
QY 3 VTSKTN 8

RESULT 2  
ID W07127 standard; peptide; 9 AA.  
AC W07127;  
DT 23-JAN-1997 (first entry)  
DE Synthetic peptide used in GalNac-transferase activity SPA.  
KW SPA; scintillation proximity assay; antigen; bead coating; capture; antibody; N-acetyl galactosamine transferase; GalNac transferase; activity; enzyme; O-linked glycosylation.  
OS Synthetic.  
PN W09615258-A1.  
PD 23-MAY-1996.  
PR 08-NOV-1995; U13483.  
PR 16-NOV-1994; US-340283.  
PA (UPJO ) UPJOHN CO.  
PI Elhammer AP;  
DR WPI; 96-26820/27.  
PT Scintillation proximity assay for N-acetyl:galactosaminyl activity - esp. for large scale screening of cpds. for their effect on enzyme activity  
PS Claim 14; Page 17; 29pp; English.  
CC W06985-W07180 are antigenic peptides derived from either the product of the human c-myc oncogene or the FLAG peptide (DYKDDK). The peptides are useful for coating beads used in a scintillation proximity assay for N-acetyl:galactosamine (GalNac)-transferase (GNT) activity. The assay involves fewer steps than known assays and is quicker, producing excellent signal-to-noise ratios. The assay is capable of screening large numbers of cpds. for their ability to affect GNT activity and is thus useful for identifying inhibitors and promoters of glycosylation (in partic. O-linked glycosylation).  
SQ Sequence 9 AA;

Query Match 17.2%; Score 5; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 tvtsk 7  
|||||  
QY 2 TVTSK 6

RESULT 3  
ID R03909 standard; peptide; 10 AA.  
AC R03909;  
DT 16-FEB-1993 (first entry)  
DE HIV-antibody reactive peptide.  
KW HIV; diagnosis.  
OS Synthetic.  
PN EP-362927-A.  
PD 11-APR-1990.  
PF 25-SEP-1989; 202398.  
PR 06-OCT-1988; NL-002447.  
PA (ALKU ) AKZO NV.  
PI Goudsmit J, Meloen RH;  
DR WPI; 90-109281/15.  
PT Nine new synthetic oligopeptide cpds. reactive with HIV antibodies - useful as diagnostic reagents, non infectious and safe to use  
PS Claim 17; Page 6 + Fig 9; 8pp; English.  
CC The peptides given in R03901-09 react immunochemically with antibodies directed against HIV. They are suitable for use in a diagnostic method for determining the presence of HIV or HIV-antibodies in a test fluid. In contrast to the native HIV, the peptides have the great advantage that these are of a safe non-infectious origin. They are made by standard (esp. solid phase) methods of peptide synthesis, or by recombinant DNA techniques.  
SQ Sequence 10 AA;

Query Match 17.2%; Score 5; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 irqga 7  
|||||  
Qy 24 IRQGA 28

## RESULT 4

ID R90327 standard; Peptide; 12 AA.  
AC R90327;  
DT 26-JUL-1996 (first entry)  
DE IL-8 epitope 8-24.  
KW Thioedoxin; flagellin; FLITRX; fusion protein; vector; pFLITRX; trxA;  
KW Thioedoxin-flagellin; flagellate bacteria; Caulobacter crescentus; flic;  
KW E.coli; Bacillus subtilis; peptide library; epitope mapping; vaccine;  
KW ligand-receptor interaction; antibody; diagnosis; therapy; IL-8;  
KW Interleukin-8; epitope.  
OS Synthetic.  
PN W09534664-A1.  
PD 21-DEC-1995.  
PF 02-MAY-1995; U05471.  
PR 16-JUN-1994; US-260582.  
PA (GENY ) GENETICS INST INC.  
PI Lu Z, McCoy JM;  
DR WPI: 96-049692/05.  
PT DNA encoding thioedoxin-flagellin fusion protein, opt. including a peptide library, used to detect ligand-target interactions, e.g. for epitope mapping  
PT for epitope mapping  
PS Claim 12; Page 44; 80pp; English.  
CC R90304-R90338 represent peptide epitopes of the region of interleukin-8 (IL-8) represented by R90371. The DNA encoding these sequences are individually cloned into either the thioedoxin or flagellin regions of the vector pFLITRX (see T11734). By cloning these sequences into pFLITRX a IL-8 peptide library is created. The vector contains the coding sequence for a thioedoxin-flagellin (FLITRX) fusion protein. The thioedoxin used is E.coli trxA and the flagellin is E.coli flic. This sequence is used to transform cells, which can then be used in a method for detecting a peptide interaction with a target molecule. The host cells are able to display the fusion protein on their surface and preferably flagellate (but flagellin-negative) non-motile bacteria such as Caulobacter crescentus, Bacillus subtilis or E.coli. The transformed cells can display random or designed peptide libraries (such as for IL-8), so that upon contact with a target molecule an interaction can be detected. This is particularly useful in epitope mapping and studies of ligand-receptor interactions, including potential effector molecules. The protein encoded by this sequence can be used in a vaccine. The fusion proteins produced can be used to raise antibodies for use in diagnosis, therapy, or purification.  
SQ Sequence 12 AA;

Query Match 17.2%; Score 5; DB 17; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.15e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 skfss 10  
|||||  
Qy 12 SKFPS 16

## RESULT 5

ID W24101 standard; peptide; 20 AA.  
AC W24101;  
DT 21-NOV-1997 (first entry)  
DE Canine immunoglobulin E peptide 4.  
KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.  
OS Canis familiaris.  
PN J09169795-A.  
PD 30-JUN-1997.  
PF 22-DEC-1995; 334381.  
PR 22-DEC-1995; JP-334381.  
PA (HITB ) HITACHI CHEM CO LTD.  
DR WPI: 97-389423/36.  
DR N-PSDB: T85650.  
PT Canine immunoglobulin E peptide fragment and related DNA - useful for the preparation of anti-canine immunoglobulin E antibody

PS Claim 2; Page 9; 12pp; Japanese.  
CC W24098-106 are peptide fragments containing at least 5 continuous amino acids of the partial canine immunoglobulin E (IgE) protein shown in W24097. The peptides are used for the preparation of anti-canine IgE antibody. The anti-canine IgE antibody can be used for the diagnosis of canine allergies.  
SQ Sequence 20 AA;

Query Match 17.2%; Score 5; DB 24; Length 20;

Best Local Similarity 100.0%; Pred. No. 4.15e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 itvts 5  
|||||  
Qy 1 ITVTS 5

## RESULT 6

ID W24100 standard; peptide; 20 AA.  
AC W24100;  
DT 21-NOV-1997 (first entry)  
DE Canine immunoglobulin E peptide 3.  
KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.  
OS Canis familiaris.  
PN J09169795-A.  
PD 30-JUN-1997.  
PF 22-DEC-1995; 334381.  
PR 22-DEC-1995; JP-334381.  
PA (HITB ) HITACHI CHEM CO LTD.  
DR WPI: 97-389423/36.  
DR N-PSDB: T85649.  
PT Canine immunoglobulin E peptide fragment and related DNA - useful for the preparation of anti-canine immunoglobulin E antibody  
PS Claim 2; Page 9; 12pp; Japanese.  
CC W24098-106 are peptide fragments containing at least 5 continuous amino acids of the partial canine immunoglobulin E (IgE) protein shown in W24097. The peptides are used for the preparation of anti-canine IgE antibody. The anti-canine IgE antibody can be used for the diagnosis of canine allergies.  
SQ Sequence 20 AA;

Query Match 17.2%; Score 5; DB 24; Length 20;

Best Local Similarity 100.0%; Pred. No. 4.15e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 itvts 15  
|||||  
Qy 1 ITVTS 5

## RESULT 7

ID R52494 standard; Peptide; 26 AA.  
AC R52494;  
DT 30-SEP-1996 (first entry)  
DE Human heavy chain surface patch HV1CSHUMAN.  
KW antibody; humanised; murine; human; heavy chain; light; variable; framework region; complementarity determining region; reshaping; modelling; surface residue; modify.  
OS Homo sapiens.  
PN EP-592106-A1.  
PD 13-APR-1994.  
PF 07-SEP-1993; 307051.  
PR 09-SEP-1992; US-942245.  
PA (PEDE/) PEDERSEN J T.  
PA (IMMU-) IMMUNOGEN INC.  
PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;  
DR WPI: 94-120230/15.  
PT Method of resurfacing of rodent antibodies to produce humanised antibody forms - for producing non-human antibodies with improved therapeutic efficiency by presenting human surface on V-region  
PS Example 1; Page 21; 230pp; English.  
CC Modification of a rodent antibody or fragment by resurfacing in order to produce a humanised rodent antibody can be determined by calculating

CC homology between murine and human antibody surfaces. In order to test  
 CC the resurfacing approach of the invention, three humanisation experiments  
 CC were set up. (1) traditional loop grafting; (2) resurfacing approach  
 CC using most similar chain; and (3) resurfacing approach using human  
 CC sequences with most similar surface residues. R52448-520 are the surface  
 CC residue patterns in human heavy chain antibody variable regions. These  
 CC "patches" were used in the third method, where rodent light and heavy  
 CC chains were matched and the most similar human sequence found  
 CC independently only over the surface residues indicated in R52030-67.  
 SQ Sequence 26 AA;

Query Match 17.2%; Score 5; DB 17; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 rqqas 21  
 |||||  
 Qy 25 RQGAS 29

RESULT 8 R52435 standard; Peptide; 26 AA.

AC R52435; 1996 (first entry)  
 DT Mouse heavy chain surface patch MUSIGHAPX.  
 DE antibody; humanised; murine; human; heavy chain; light; variable;  
 KW framework region; complementarity determining region; reshaping;  
 KW modelling; surface residue; modify.  
 OS MUS SP.  
 PN EP-592106-A1.  
 PD 13-APR-1994.  
 PF 07-SEP-1993; 307051.  
 PR 09-SEP-1992; US-942245.  
 PA (PEDE/) PEDERSEN J T.  
 PA (IMMU-) IMMUNOGEN INC.  
 PI Guild BC, Pedersen JT, Rees AR, Roguska MA, Searle SMJ;  
 DR WPI; 94-120230/15.  
 PT Method of resurfacing of rodent antibodies to produce humanised  
 PT antibody forms - for producing non-human antibodies with improved  
 PT therapeutic efficiency by presenting human surface on V-region  
 PS Example 1; Page 20; 230pp; English.  
 CC Modification of a rodent antibody or fragment by resurfacing in order  
 CC to produce a humanised rodent antibody can be determined by calculating  
 CC homology between murine and human antibody surfaces. In order to test  
 CC the resurfacing approach of the invention, three humanisation experiments  
 CC were set up. (1) traditional loop grafting; (2) resurfacing approach  
 CC using most similar chain; and (3) resurfacing approach using human  
 CC sequences with most similar surface residues. R52243-447 are the surface  
 CC residue patterns in mouse heavy chain antibody variable regions. These  
 CC "patches" were used in the third method, where rodent light and heavy  
 CC chains were matched and the most similar human sequence found  
 CC independently only over the surface residues indicated in R52030-67.  
 SQ Sequence 26 AA;

Query Match 17.2%; Score 5; DB 17; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 20 tsktn 24  
 |||||  
 Qy 4 TSKTN 8

RESULT 9 R7217 standard; Peptide; 32 AA.

AC R7217;  
 DT 23-AUG-1995 (first entry)  
 DE Mouse anti-human IL-6 Ab H chain V region framework region 3.  
 KW Primer; PCR; amplify; kappa; light chain; variable region; mouse; human;  
 KW interleukin; antibody; hybridoma; CDR; framework; constant region;  
 KW heavy chain; disorder; antigenicity.  
 OS Synthetic.  
 PN WO9428159-A.

PD 08-DEC-1994.  
 PF 30-MAY-1994; J00859.  
 PR 31-MAY-1993; JP-129787.  
 PA (CHUS ) CHUGAI SEIVAKU KK.  
 PA (CHUS ) CHUGAI PHARM CO LTD.  
 PI Hirata Y, Sato K, Tsuchiya M;  
 DR WPI; 95-022828/03.  
 PT Antibody against IL-6 - useful for the therapy and treatment of  
 PT IL-6 related disorders.  
 PS Claim 13; Page 66; 82pp; Japanese.  
 CC The sequence of the mouse anti-human interleukin-6 (IL-6) antibody heavy  
 CC chain variable region framework region (FR) 3. The sequences of FR1-4  
 CC regions 1-3 (R7212-4) were used in conjunction with the complementarity determining  
 CC regions 1-3 (R7212-4) to construct a chimaeric antibody against human  
 CC interleukin-6 (IL-6). The vectors Q75914-7 express constructs encoding  
 CC fragments of a chimaeric antibody to the human IL-6 comprising (a) a  
 CC light chain with (1) a variable region containing 3 CDR (R7201-3)  
 CC inserted into several framework regions (FR) (R7204-7) and (ii) a human  
 CC light chain constant region and (b) a heavy chain with (i) a variable  
 CC region containing 3 CDR (R7212-4) inserted into FR (R7215-8) and (ii)  
 CC a human light chain constant region. The FR of the light chain may be  
 CC mouse derived (Q75888) or from the human antibody REI. The heavy chain  
 CC FR may also be mouse derived (Q75889) or from the human antibody DAW.  
 CC The antibodies can be used in the treatment of IL-6 related disorders.  
 CC The antibodies are useful as they have low antigenicity due to the use of  
 CC human derived sequences and low antigenicity mouse derived sequences.  
 SQ Sequence 32 AA;

Query Match 17.2%; Score 5; DB 12; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 kdtsk 10  
 |||||  
 Qy 9 KDTSK 13

RESULT 10 R11409 standard; Protein; 43 AA.  
 AC R11409;  
 DT 10-JUN-1991 (first entry)  
 DE Hepatitis B surface antigen pps1-110-148 complex epitope.  
 KW HBsAg; chimaera; hepadnavirus; vaccine; pPV-Nhe.  
 OS Synthetic.  
 PN EP-421635-A.  
 PD 10-APR-1991.  
 PF 19-SEP-1990; 310264.  
 PR 19-SEP-1989; GB-021171.  
 PR 13-AUG-1990; GB-017728.  
 PA (WELL ) WELLCOME FOUNDATION LTD.  
 PI Brown AL, Clarke BE, Rowlands DJ;  
 DR WPI; 91-103861/15.  
 DR N-PSDB; Q11233.

PT Particles comprising chimeric hepadnavirus core antigen protein -  
 PT contg. foreign aminoacid sequence comprising epitope of eg  
 PT hepatitis-A or -B and human immuno-deficiency virus, used as vaccine.  
 PS Example 3; Page 10; 24pp; English.  
 CC The insert encoding this epitope was ligated into the NheI restriction  
 CC site of plasmid pPV-Nhe to recreate the NheI sites which code for  
 CC Ala-Ser residues flanking the epitope. The recombinant plasmid was  
 CC transformed into E.coli strain XL-1 Blue. Positive clones were cultured  
 CC to high density in nutrient broth and induced to express chimaeric  
 CC protein comprising the desired epitope by addition of IPTG.  
 SQ Sequence 43 AA;

Query Match 17.2%; Score 5; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 4.15e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 25 skfps 29  
 |||||  
 Qy 12 SKFPS 16

RESULT 11  
 ID R11482 standard; Protein; 70 AA.  
 AC R11482;  
 DE Membrane anchoring peptide encoded by exons I and II of human IgG  
 DE gamma 4 heavy chain.  
 DE gamma 4 heavy chain.  
 KW Autoimmune diseases; migis epitope; rheumatoid arthritis;  
 KW multiple sclerosis; systemic lupus erythematosus.  
 FH Key  
 FT peptide  
 FT Location/Qualifiers  
 FT 1..18  
 FT /label= extracellular segment  
 PN WO9104055-A.  
 PD 04-APR-1991.  
 PF 14-SEP-1990; U05229.  
 PR 15-SEP-1989; US-408123.  
 PA (TANO-) TANOX BIOSYSTEMS IN.  
 PI Chang TW;  
 DR WPI: 91-117334/16.  
 DR N-PSDB; Q11351, Q11352.  
 PT Treating auto-immune diseases associated with partic.  
 PT immunoglobulin - by administering peptide corres. to MIGIS-mu or  
 PT MIGIS-gamma segments  
 PS Disclosure: fig 2: 44pp; English.  
 CC This sequence is the extracellular membrane anchoring peptide  
 CC segment of the IgG gamma 4 heavy chain. It contains the migis-  
 CC gamma epitope at its N-terminal and can be used as an immunogen  
 CC to raise anti-migis antibodies which can be used to suppress or  
 CC deplete B-cells or precursors.  
 CC See also Q11345-50, Q11353 and R11483.  
 CC Sequence 70 AA;  
 SQ

Query Match 17.2%; Score 5; DB 2; Length 70;  
 Best Local Similarity 100.0%; Pred. No. 4.15e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 66 irqga 70  
 |||||  
 QY 24 IRQGA 28

RESULT 12  
 ID R11480 standard; Protein; 71 AA.  
 AC R11480;  
 DE Membrane anchoring peptide encoded by exons I and II of human IgG  
 DE gamma 2 heavy chain.  
 DE gamma 2 heavy chain.  
 KW Autoimmune diseases; migis epitope; rheumatoid arthritis;  
 KW multiple sclerosis; systemic lupus erythematosus.  
 FH Key  
 FT peptide  
 FT Location/Qualifiers  
 FT 1..18  
 FT /label= extracellular segment  
 PN WO9104055-A.  
 PD 04-APR-1991.  
 PF 14-SEP-1990; U05229.  
 PR 15-SEP-1989; US-408123.  
 PA (TANO-) TANOX BIOSYSTEMS IN.  
 PI Chang TW;  
 DR WPI: 91-117334/16.  
 DR N-PSDB; Q11347, Q11348.  
 PT Treating auto-immune diseases associated with partic.  
 PT immunoglobulin - by administering peptide corres. to MIGIS-mu or  
 PT MIGIS-gamma segments  
 PS Disclosure: fig 2: 44pp; English.  
 CC This sequence is the extracellular membrane anchoring peptide  
 CC segment of the IgG gamma 2 heavy chain. It contains the migis-  
 CC gamma epitope at its N-terminal and can be used as an immunogen  
 CC to raise anti-migis antibodies which can be used to suppress or  
 CC deplete B-cells or precursors.  
 CC See also Q11345-46, Q11349-53 and R11483.  
 CC Sequence 71 AA;  
 SQ

Query Match 17.2%; Score 5; DB 2; Length 71;

Best Local Similarity 100.0%; Pred. No. 4.15e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 irqga 71  
 |||||  
 QY 24 IRQGA 28

RESULT 13  
 ID R47919 standard; Protein; 71 AA.  
 AC R47919;  
 DE Segment of human gamma-2 chain.  
 DE 20-JUL-1994 (first entry)  
 DE Isoform: peptide; anchoring protein; membrane spanning protein;  
 KW Isoform: peptide; anchoring protein; membrane spanning protein;  
 KW transmembrane domain; immunoglobulin; B cell; immune system.  
 OS Homo sapiens.  
 FH Key  
 FT region  
 FT 1..18  
 FT Location/Qualifiers  
 FT /label= Extracellular segment.  
 PN US5281699-A.  
 PD 25-JAN-1994.  
 PF 01-JUN-1990; 531787.  
 PR 01-JUN-1990; US-531787.  
 PA (TANO-) TANOX BIOSYSTEMS INC.  
 PI Chang TW;  
 DR WPI: 94-042862/05.  
 DR N-PSDB; Q55785, Q66079.  
 DR New peptide(s) representing extracellular immunoglobulin gamma  
 PT epitope(s) - for generating isotype specific antibodies, used to  
 PT kill B cells involved in leukaemia and lymphoma  
 PS Disclosure: Figure 4B; 22pp; English.  
 CC The heavy chains of immunoglobulins contain membrane anchoring  
 CC peptides which span the cell membrane lipid bilayer, fixing the  
 CC associated immunoglobulin to the cell membrane surface. The  
 CC extracellular portions of these peptides are unique for different  
 CC isotypes, but tend to be similar among different subclasses of a  
 CC particular isotype. The extracellular segment forms an epitope  
 CC unique to the B cells which produce each isotype. These membrane  
 CC bound immunoglobulin isotype-specific extracellular epitopes are  
 CC not present on the secreted, soluble form of the immunoglobulins.  
 CC Peptides from the extracellular segment of human gamma chain can be  
 CC used to produce antibodies which, when coupled to a cytotoxic  
 CC compound, can kill B cells which produce IgG (but not those  
 CC producing immunoglobulins of other isotypes). Only cells bearing a  
 CC particular immunoglobulin isotype will be destroyed by antibody,  
 CC leaving other components of the immune system intact. This peptide  
 CC sequence is the anchoring peptide of the human gamma-2 chain.  
 CC Sequence 71 AA;  
 SQ

Query Match 17.2%; Score 5; DB 9; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 4.15e-02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 irqga 71  
 |||||  
 QY 24 IRQGA 28

RESULT 14  
 ID R47921 standard; Protein; 71 AA.  
 AC R47921;  
 DE Segment of human gamma-4 chain.  
 DE 20-JUL-1994 (first entry)  
 DE Isoform: peptide; anchoring protein; membrane spanning protein;  
 KW Isoform: peptide; anchoring protein; membrane spanning protein;  
 KW transmembrane domain; immunoglobulin; B cell; immune system.  
 OS Homo sapiens.  
 FH Key  
 FT region  
 FT 1..18  
 FT Location/Qualifiers  
 FT /label= Extracellular segment.  
 PN US5281699-A.  
 PD 25-JAN-1994.  
 PF 01-JUN-1990; 531787.  
 PR 01-JUN-1990; US-531787.

PA (TANO-) TANOX BIOSYSTEMS INC.

PI Chang TW;

DR WPI: 94-042862/05.

DR N-PSDB: Q55787, Q66081.

PT New peptide(s), representing extracellular immunoglobulin gamma

PT epitope(s) - for generating isotype specific antibodies, used to

PT kill B cells involved in leukemia and lymphoma

PS Disclosure; Figure 4B; 22pp; English.

CC The heavy chains of immunoglobulins contain membrane anchoring

CC peptides which span the cell membrane lipid bilayer, fixing the

CC associated immunoglobulin to the cell membrane surface. The

CC extracellular portions of these peptides are unique for different

CC isotypes, but tend to be similar among different subclasses of a

CC particular isotype. The extracellular segment forms an epitope

CC unique to the B cells which produce each isotype. These membrane

CC bound immunoglobulin isotype-specific extracellular epitopes are

CC not present on the secreted, soluble form of the immunoglobulins.

CC Peptides from the extracellular segment of human gamma chain can be

CC used to produce antibodies which, when coupled to a cytotoxic

CC compound, can kill B cells which produce IgG (but not those

CC producing immunoglobulins of other isotypes). Only cells bearing a

CC particular immunoglobulin isotype will be destroyed by antibody,

CC leaving other components of the immune system intact. This peptide

CC sequence is the anchoring peptide of the human gamma-4 chain.

SQ Sequence 71 AA;

Query Match 17.2%; Score 5; DB 9; Length 71;

Best Local Similarity 100.0%; Pred. No. 4.15e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 irqga 71

Qy 24 IROGA 28

RESULT 15

ID P91702 standard; protein: 79 AA.

AC P91702;

DT 13-JUN-1990 (first entry)

DE Protein increasing pulmonary surfactant activity.

KW Pulmonary surfactant; respiratory disorders.

OS Homo sapiens.

PN W08900167-A.

PD 12-JAN-1989.

PF 29-JUN-1988; 00361.

PR 01-JUL-1987; SE-027249.

PR 22-SEP-1987; SE-036612.

PA (KABI) Kabigen Ab.

PI Curstedt T, Robertsson B, Jornvall H;

DR WPI: 89-039631/05.

PT Proteins with pulmonary surfactant activity -

PT obtd. from pig lung and human broncho-alveolar lavage or

PT amniotic fluid, for treating respiratory disorders.

PS Claim 5; Page 16; 24pp; English.

CC Proteins, derived from bronchoalveolar lavage and amniotic fluid, can be

CC extracted and shown to have pulmonary surfactant activity. Useful in

CC treating respiratory disorders, reducing surface tension at air-liquid

CC interface.

SQ Sequence 79 AA;

Query Match 17.2%; Score 5; DB 1; Length 79;

Best Local Similarity 100.0%; Pred. No. 4.15e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 38 splvv 42

Qy 16 SPLVV 20

Search completed: Sat Aug 28 15:00:28 1999

Job time : 76 secs.



\*\*\*\*\*  
WQSEH  
\*\*\*\*\* (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:58:06 1999; MasPar time 5.01 Seconds  
231.854 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-46  
Description: (1-29) from US09049695.pep  
Perfect Score: 29  
Sequence: 1 ITVTSKTNKDTSKFSPPLVYVNIROGAS 29

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 122810 seqs, 40068593 residues  
Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 2.709; Variance 0.427; scale 6.345

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result	No.	Score	Query	Match	Length	ID	Description	Pred.	No.
1	7	24.1	612	2	S29318	lysophospholipase (EC 1.29e-01			
2	6	20.7	152	2	D73603	hypothetical protein	6.61e+00		
3	6	20.7	205	2	D69134	conserved hypothetical	6.61e+00		
4	6	20.7	279	2	A24469	beta-lactamase (EC 3.6.1e+00			
5	6	20.7	321	2	E69687	cytochrome aa3 quinol	6.61e+00		
6	6	20.7	362	2	S50993	ribosomal protein L4.	6.61e+00		
7	6	20.7	362	2	S45887	ribosomal protein L4.	6.61e+00		
8	6	20.7	385	2	S61566	BBP1 protein - yeast	6.61e+00		
9	6	20.7	457	2	A25351	ubiquinol--cytochrome	6.61e+00		
10	6	20.7	486	1	FOLJEV	gag polyprotein - equ	6.61e+00		
11	6	20.7	498	2	T7411	integral membrane pro	6.61e+00		
12	6	20.7	583	2	T02382	hypothetical protein	6.61e+00		
13	6	20.7	603	2	S47917	rab geranylgeranyl tr	6.61e+00		
14	6	20.7	660	2	B34469	pullulanase secretion	6.61e+00		
15	6	20.7	913	3	JG0168	gob-5 protein - Mouse	6.61e+00		
16	6	20.7	1140	2	T37786	hypothetical protein	6.61e+00		
17	6	20.7	1291	3	T00019	period protein homolo	6.61e+00		
18	6	20.7	1567	2	T03730	antigen containing ep	6.61e+00		
19	6	20.7	1857	2	S01787	fatty-acid synthase (	6.61e+00		
20	5	17.2	10	2	C39111	Ig heavy chain C regi	2.25e+02		
21	5	17.2	26	2	I45087	cysteine proteinase h	2.25e+02		
22	5	17.2	41	1	S00690	cytochrome b559 compo	2.25e+02		
23	5	17.2	53	2	B36005	Ig heavy chain V regi	2.25e+02		

24	5	17.2	56	2	PD0016	dextranucrase (EC 2.2.25e+02			
25	5	17.2	56	2	S01277	NADH dehydrogenase (u	2.25e+02		
26	5	17.2	67	4	S07693	hypothetical protein	2.25e+02		
27	5	17.2	71	2	JQ2195	hypothetical 8.1k pro	2.25e+02		
28	5	17.2	74	2	S26793	Ig heavy chain V regi	2.25e+02		
29	5	17.2	78	2	S46461	Ig heavy chain V regi	2.25e+02		
30	5	17.2	85	2	I64137	phosphotransferase sy	2.25e+02		
31	5	17.2	85	1	WQECPH	phosphotransferase sy	2.25e+02		
32	5	17.2	85	1	S12749	phosphotransferase sy	2.25e+02		
33	5	17.2	85	1	WQEBPH	phosphotransferase sy	2.25e+02		
34	5	17.2	91	2	S54546	hypothetical protein	2.25e+02		
35	5	17.2	93	2	G30010	hypothetical ORF-8 pr	2.25e+02		
36	5	17.2	93	2	S14020	hypothetical protein	2.25e+02		
37	5	17.2	93	2	I48636	rearranged T-cell rec	2.25e+02		
38	5	17.2	94	2	A70540	hypothetical protein	2.25e+02		
39	5	17.2	96	2	S26922	Ig heavy chain V regi	2.25e+02		
40	5	17.2	96	2	S26924	Ig heavy chain V regi	2.25e+02		
41	5	17.2	96	2	S26923	Ig heavy chain V regi	2.25e+02		
42	5	17.2	97	2	D64600	conserved hypothetical	2.25e+02		
43	5	17.2	98	2	S26159	NADH dehydrogenase (u	2.25e+02		
44	5	17.2	98	2	S41843	NADH dehydrogenase (u	2.25e+02		
45	5	17.2	98	2	H71630	hypothetical protein	2.25e+02		
46	5	17.2	102	2	S62341	L71-9 protein - fruit	2.25e+02		
47	5	17.2	103	2	A71218	hypothetical protein	2.25e+02		
48	5	17.2	104	2	S54132	hypothetical protein	2.25e+02		
49	5	17.2	104	2	S50911	metallothionein-2 - T	2.25e+02		
50	5	17.2	105	1	RIIDS2	somatostatin-22 precu	2.25e+02		
51	5	17.2	108	1	WZBE1	gene 1 protein - huma	2.25e+02		
52	5	17.2	110	2	A71451	hypothetical protein	2.25e+02		
53	5	17.2	113	2	S24247	Ig heavy chain V regi	2.25e+02		
54	5	17.2	118	2	S18556	Ig heavy chain V regi	2.25e+02		
55	5	17.2	119	1	GIHUDW	Ig heavy chain V-II r	2.25e+02		
56	5	17.2	119	2	S18555	Ig heavy chain V regi	2.25e+02		
57	5	17.2	121	2	A36005	Ig heavy chain V regi	2.25e+02		
58	5	17.2	123	2	T01778	hypothetical protein	2.25e+02		
59	5	17.2	124	2	A49002	Ig heavy chain V regi	2.25e+02		
60	5	17.2	125	2	A31557	T-cell receptor alpha	2.25e+02		
61	5	17.2	125	2	JU0179	heparin-binding prote	2.25e+02		
62	5	17.2	128	2	S52084	ribosomal protein L22	2.25e+02		
63	5	17.2	128	2	JC2119	heparin-binding prote	2.25e+02		
64	5	17.2	128	2	JC2121	heparin-binding prote	2.25e+02		
65	5	17.2	128	2	JC2120	heparin-binding prote	2.25e+02		
66	5	17.2	131	2	S77285	hypothetical protein	2.25e+02		
67	5	17.2	132	2	A23706	amine dehydrogenase (	2.25e+02		
68	5	17.2	134	2	B64062	opacity-associated pr	2.25e+02		
69	5	17.2	137	2	JC4233	lysozyme (EC 3.2.1.17	2.25e+02		
70	5	17.2	138	2	I46634	rearranged T-cell rec	2.25e+02		
71	5	17.2	145	2	I46633	rearranged T-cell rec	2.25e+02		
72	5	17.2	147	1	G2HUCS	Ig heavy chain precu	2.25e+02		
73	5	17.2	150	2	C64337	hypothetical protein	2.25e+02		
74	5	17.2	151	2	S43727	hypothetical protein	2.25e+02		
75	5	17.2	153	2	S58296	hypothetical protein	2.25e+02		
76	5	17.2	153	2	T01831	hypothetical protein	2.25e+02		
77	5	17.2	164	2	H64167	hypothetical protein	2.25e+02		
78	5	17.2	165	2	H69851	hypothetical protein	2.25e+02		
79	5	17.2	168	2	S60970	hypothetical protein	2.25e+02		
80	5	17.2	176	2	JC6152	orphanin FQ precursor	2.25e+02		
81	5	17.2	176	2	S76001	phosphoribosylaminol	2.25e+02		
82	5	17.2	178	2	E70069	conserved hypothetical	2.25e+02		
83	5	17.2	178	2	S77880	probable GTP-binding	2.25e+02		
84	5	17.2	179	2	UQ0345	ubiquinol--cytochrome	2.25e+02		
85	5	17.2	181	1	RKMUB3	ribulose-bisphosphate	2.25e+02		
86	5	17.2	181	2	S20528	hypothetical protein	2.25e+02		
87	5	17.2	181	1	24BPT9	gene 55.13 protein -	2.25e+02		
88	5	17.2	182	2	A70961	hypothetical protein	2.25e+02		
89	5	17.2	187	2	F69993	carbonic anhydrase ho	2.25e+02		
90	5	17.2	188	2	S75643	hypothetical protein	2.25e+02		
91	5	17.2	191	2	S06618	SPR6 protein - yeast	2.25e+02		
92	5	17.2	199	2	H70756	hypothetical protein	2.25e+02		
93	5	17.2	202	2	B45512	cold-regulated protei	2.25e+02		
94	5	17.2	207	2	S60669	superoxide dismutase	2.25e+02		
95	5	17.2	207	2	A56190	titin - rat (fragment	2.25e+02		
96	5	17.2	209	2	F69839	conserved hypothetical	2.25e+02		

97	5	17.2	210	2	S77737	opacity protein opad	2.25e+02	170	5	17.2	314	2	JQ2363	polyprotein - bean co	2.25e+02
98	5	17.2	210	2	S77469	hypothetical protein	2.25e+02	171	5	17.2	314	2	JQ2362	polyprotein - bean co	2.25e+02
99	5	17.2	215	2	B57055	melanocortin-4 recept	2.25e+02	172	5	17.2	314	2	S56055	hypothetical protein	2.25e+02
100	5	17.2	215	2	D71634	hypothetical protein	2.25e+02	173	5	17.2	315	2	S75143	sensory transduction	2.25e+02
101	5	17.2	216	1	TUBPX2	tail fiber protein gp	2.25e+02	174	5	17.2	317	2	G69535	carotenoid biosynthet	2.25e+02
102	5	17.2	217	2	D33216	ependymal precursor -	2.25e+02	175	5	17.2	318	2	D64815	ybHN protein - Escher	2.25e+02
103	5	17.2	217	2	D33216	prostaglandin E2 rece	2.25e+02	176	5	17.2	318	2	T00552	lyso-phospholipase hom	2.25e+02
104	5	17.2	218	2	S58769	brain-specific protei	2.25e+02	177	5	17.2	319	2	S40175	ExoW protein - Rhizob	2.25e+02
105	5	17.2	218	2	S56556	hypothetical 24.6K pr	2.25e+02	178	5	17.2	323	2	G64240	hypothetical protein	2.25e+02
106	5	17.2	226	2	S33280	homeotic protein msh-	2.25e+02	179	5	17.2	323	2	S05035	protein kinase (EC 2.	2.25e+02
107	5	17.2	226	2	S14006	hypothetical protein,	2.25e+02	180	5	17.2	325	2	S27796	hypothetical protein	2.25e+02
108	5	17.2	234	2	S14469	asparagine-rich prote	2.25e+02	181	5	17.2	325	2	S42832	F40F12.3 protein - Ca	2.25e+02
109	5	17.2	240	2	T02018	heat shock protein 26	2.25e+02	182	5	17.2	327	2	T01478	hypothetical protein	2.25e+02
110	5	17.2	241	2	S16004	heat shock protein 21	2.25e+02	183	5	17.2	328	2	S67592	hypothetical protein	2.25e+02
111	5	17.2	244	2	S73539	triosphosphate isome	2.25e+02	184	5	17.2	328	2	S15284	isopenicillin N synth	2.25e+02
112	5	17.2	245	2	H70485	hypothetical protein	2.25e+02	185	5	17.2	332	2	A70257	immunogenic protein P	2.25e+02
113	5	17.2	245	1	CHUQC	complement subcompone	2.25e+02	186	5	17.2	332	2	A57055	melanocortin receptor	2.25e+02
114	5	17.2	250	2	S38888	COI intron 12 protein	2.25e+02	187	5	17.2	336	2	S152585	lymphocyte-specific p	2.25e+02
115	5	17.2	251	2	E48227	3-isopropylmalate deh	2.25e+02	188	5	17.2	336	2	G69270	flavoprotein (fprA-1)	2.25e+02
116	5	17.2	252	2	B71105	hypothetical protein	2.25e+02	189	5	17.2	337	2	H64389	isocitrate dehydrogen	2.25e+02
117	5	17.2	252	2	S05629	cytochrome-c oxidase	2.25e+02	190	5	17.2	339	2	A43512	lymphocyte-specific p	2.25e+02
118	5	17.2	253	2	A38201	transcription factor	2.25e+02	191	5	17.2	339	2	A43542	lymphocyte antigen Wp	2.25e+02
119	5	17.2	255	2	H70231	probable stage 0 spor	2.25e+02	192	5	17.2	341	1	QXFF21	NADH dehydrogenase (u	2.25e+02
120	5	17.2	255	2	G18834	iron (III) dicitrate	2.25e+02	193	5	17.2	342	2	S55675	Gal-beta-1,3GalNac al	2.25e+02
121	5	17.2	255	2	H64630	iron(III) dicitrate A	2.25e+02	194	5	17.2	343	2	S77079	hypothetical protein	2.25e+02
122	5	17.2	258	2	S77158	hydrogenase accessory	2.25e+02	195	5	17.2	344	2	C70401	phospho-N-acetylmuram	2.25e+02
123	5	17.2	258	2	S08514	opacity protein-relat	2.25e+02	196	5	17.2	344	2	D70883	probable oxidoreducta	2.25e+02
124	5	17.2	261	2	A60914	per region circadian	2.25e+02	197	5	17.2	345	2	S73729	MG307 homolog H08.Orf	2.25e+02
125	5	17.2	262	2	B42478	glutamine-binding pro	2.25e+02	198	5	17.2	347	1	SYEQQA	quinolinate synthetas	2.25e+02
126	5	17.2	262	2	S18130	alpha-amylase inhibit	2.25e+02	199	5	17.2	348	2	F70179	spemidine/putrescine	2.25e+02
127	5	17.2	266	2	JC1071	coat protein - soybea	2.25e+02	200	5	17.2	353	2	A43633	41K protein - Lactoba	2.25e+02
128	5	17.2	266	2	T01326	alpha zein 8, 22K - m	2.25e+02	201	5	17.2	355	1	QQCVW1	AV1 protein - abutilo	2.25e+02
129	5	17.2	267	2	S18931	coat protein - soybea	2.25e+02	202	5	17.2	357	2	S75954	hypothetical protein	2.25e+02
130	5	17.2	269	2	S75243	hypothetical protein	2.25e+02	203	5	17.2	363	2	S38154	hypothetical protein	2.25e+02
131	5	17.2	269	2	S18115	hypothetical protein	2.25e+02	204	5	17.2	365	2	C37753	quinolinate synthase	2.25e+02
132	5	17.2	270	2	E69921	DNA ligase homolog yo	2.25e+02	205	5	17.2	370	2	A48340	capsid protein - pean	2.25e+02
133	5	17.2	271	2	J02285	nodulin-26 - soybean	2.25e+02	206	5	17.2	370	2	S15013	wnt-1 protein - zebra	2.25e+02
134	5	17.2	272	2	S72576	oligopeptide transpor	2.25e+02	207	5	17.2	370	2	S69718	hypothetical protein	2.25e+02
135	5	17.2	273	2	B27035	acid phosphatase (EC	2.25e+02	208	5	17.2	371	1	HUBBHA	hyaluronoglucosaminid	2.25e+02
136	5	17.2	273	2	J00418	pyroline-5-carboxyla	2.25e+02	209	5	17.2	373	2	S48226	3-isopropylmalate deh	2.25e+02
137	5	17.2	273	2	D59633	glutamine ABC transpo	2.25e+02	210	5	17.2	373	2	S48225	3-isopropylmalate deh	2.25e+02
138	5	17.2	275	2	S32410	tryptase (EC 3.4.21.5	2.25e+02	211	5	17.2	373	2	S48228	3-isopropylmalate deh	2.25e+02
139	5	17.2	276	2	F69307	conserved hypotheticala	2.25e+02	212	5	17.2	374	2	S69339	Ig heavy chain V regi	2.25e+02
140	5	17.2	279	2	S64771	probable membrane pro	2.25e+02	213	5	17.2	374	2	S72664	Ig heavy chain V regi	2.25e+02
141	5	17.2	280	2	S71717	rRNA methylase - Stre	2.25e+02	214	5	17.2	374	2	D24827	heat shock 82K protei	2.25e+02
142	5	17.2	281	2	H71227	hypothetical protein	2.25e+02	215	5	17.2	374	2	T00619	hypothetical protein	2.25e+02
143	5	17.2	282	2	S46793	vacuolar protein sort	2.25e+02	216	5	17.2	375	2	B24827	heat shock 82K protei	2.25e+02
144	5	17.2	283	2	S27859	piroplasm surface pro	2.25e+02	217	5	17.2	375	2	C24827	heat shock 82K protei	2.25e+02
145	5	17.2	283	2	A48460	piroplasm surface pro	2.25e+02	218	5	17.2	375	1	KIBET3	thymidine kinase (EC	2.25e+02
146	5	17.2	283	2	S27863	Ts-32K protein - Thei	2.25e+02	219	5	17.2	377	2	B70140	hypothetical protein	2.25e+02
147	5	17.2	283	2	E49348	succinoglycan biosynt	2.25e+02	220	5	17.2	377	2	I50147	3-beta-hydroxy-Delta5-	2.25e+02
148	5	17.2	285	2	B44886	yadU protein - Escher	2.25e+02	221	5	17.2	378	2	JC5649	N-acetylglucosamine-6	2.25e+02
149	5	17.2	286	2	S07193	chorion protein s36 -	2.25e+02	222	5	17.2	379	2	S17979	pectin lyase (EC 4.2.	2.25e+02
150	5	17.2	287	2	S25556	genome polyprotein -	2.25e+02	223	5	17.2	381	2	S45766	hypothetical protein	2.25e+02
151	5	17.2	288	2	G70256	immunogenic protein P	2.25e+02	224	5	17.2	382	2	B64209	prolipo-protein diacyl	2.25e+02
152	5	17.2	293	2	S09209	chorion protein s36 -	2.25e+02	225	5	17.2	382	2	S52585	dycopene cyclase - Er	2.25e+02
153	5	17.2	299	2	G71088	hypothetical protein	2.25e+02	226	5	17.2	382	2	B49193	type II activin recep	2.25e+02
154	5	17.2	302	2	PS0084	genome polyprotein -	2.25e+02	227	5	17.2	382	2	A37018	N-acetylglucosamine-6	2.25e+02
155	5	17.2	303	2	PS0081	genome polyprotein -	2.25e+02	228	5	17.2	386	2	I59387	synaptotagmin V - rat	2.25e+02
156	5	17.2	304	2	E70677	hypothetical protein	2.25e+02	229	5	17.2	388	2	H71115	hypothetical protein	2.25e+02
157	5	17.2	305	1	WWBEP5	ribonucleoside-diphos	2.25e+02	230	5	17.2	389	1	SAVLVE	large surface antigen	2.25e+02
158	5	17.2	305	2	I57039	genomic screen homeob	2.25e+02	231	5	17.2	390	2	G71966	molybdopterin biosynt	2.25e+02
159	5	17.2	305	2	A40573	clathrin heavy chain	2.25e+02	232	5	17.2	391	2	D64541	molybdopterin biosynt	2.25e+02
160	5	17.2	305	2	S55655	ribonucleoside-diphos	2.25e+02	233	5	17.2	392	2	H71526	probable heat shock p	2.25e+02
161	5	17.2	307	2	G63430	quinolinate synthetas	2.25e+02	234	5	17.2	392	2	S69788	prGH protein - salmon	2.25e+02
162	5	17.2	307	2	G69505	hypothetical protein	2.25e+02	235	5	17.2	396	2	A48336	coat protein - South	2.25e+02
163	5	17.2	308	2	H65014	hypothetical protein	2.25e+02	236	5	17.2	397	2	JE0082	GPI-linked receptor p	2.25e+02
164	5	17.2	309	2	I51900	carbonic anhydrase IV	2.25e+02	237	5	17.2	399	2	D69863	hypothetical protein	2.25e+02
165	5	17.2	310	3	JC6502	pectate lyase (EC 4.2	2.25e+02	238	5	17.2	400	2	S64729	protein secretion pro	2.25e+02
166	5	17.2	310	2	E64751	probable membrane pro	2.25e+02	239	5	17.2	403	2	T01810	reverse transcriptase	2.25e+02
167	5	17.2	313	2	T02534	hypothetical protein	2.25e+02	240	5	17.2	406	2	A48059	oncoprotein zc-Myc -	2.25e+02
168	5	17.2	314	2	S75237	hypothetical protein	2.25e+02	241	5	17.2	407	1	RSFFL1	ribosomal protein DL1	2.25e+02
169	5	17.2	314	2	JQ2364	polyprotein - bean co	2.25e+02	242	5	17.2	410	2	S74951	Na+/H+-exchanging pro	2.25e+02

243	5	17.2	410	2	D70884	probable Acyl-CoA Deh	2.25e+02	316	5	17.2	523	4	D55066	probable tyrosine dec	2.25e+02
244	5	17.2	411	2	C30091	capA protein - Bacill	2.25e+02	317	5	17.2	527	2	JC5357	laccase (EC 1.10.3.2)	2.25e+02
245	5	17.2	411	1	HYSNFA	fructolase (EC 3.4.24.	2.25e+02	318	5	17.2	528	2	C69323	conserved hypothetica	2.25e+02
246	5	17.2	413	2	JE0142	glutamate synthase (E	2.25e+02	319	5	17.2	528	2	C40829	activin receptor isof	2.25e+02
247	5	17.2	417	2	B40016	matrin 3 - human (fra	2.25e+02	320	5	17.2	530	2	JC5939	estrogen receptor bet	2.25e+02
248	5	17.2	417	1	A42843	heat shock protein Hs	2.25e+02	321	5	17.2	532	2	S78176	transcription initiat	2.25e+02
249	5	17.2	417	1	A40968	heat shock protein 47	2.25e+02	322	5	17.2	536	2	A40829	activin receptor isof	2.25e+02
250	5	17.2	419	2	B69819	conserved hypothetica	2.25e+02	323	5	17.2	536	2	H71563	hypothetical protein	2.25e+02
251	5	17.2	420	2	A53531	oncofetal typhoblast	2.25e+02	324	5	17.2	538	2	S64842	probable membrane pro	2.25e+02
252	5	17.2	422	2	S24451	terminase - phage spp	2.25e+02	325	5	17.2	544	2	A71260	hypothetical protein	2.25e+02
253	5	17.2	422	2	C70565	hypothetical protein	2.25e+02	326	5	17.2	546	2	I48899	corticactin - mouse	2.25e+02
254	5	17.2	423	2	C64927	probable membrane pro	2.25e+02	327	5	17.2	547	2	E70720	probable ilvG protein	2.25e+02
255	5	17.2	424	2	C34252	isovaleryl-CoA dehydr	2.25e+02	328	5	17.2	550	2	A48063	mammary tumor/squamou	2.25e+02
256	5	17.2	425	2	A41099	protein kinase (EC 2.	2.25e+02	329	5	17.2	550	2	S65753	beta-fructofuranosida	2.25e+02
257	5	17.2	426	2	JQ1696	pisill extensin-like	2.25e+02	330	5	17.2	551	2	A57189	secY protein homolog	2.25e+02
258	5	17.2	427	2	S60742	alpha-transinducing f	2.25e+02	331	5	17.2	551	2	S53377	gZF3 protein - yeast	2.25e+02
259	5	17.2	428	2	JQ2203	UL48h protein - Marek	2.25e+02	332	5	17.2	552	2	G69464	acetolactate synthase	2.25e+02
260	5	17.2	428	2	A35534	tetrahydrofolylpolygl	2.25e+02	333	5	17.2	553	2	A49364	59 protein, brain - h	2.25e+02
261	5	17.2	429	1	S44778	transcription factor	2.25e+02	334	5	17.2	553	2	T04416	secY protein homolog	2.25e+02
262	5	17.2	434	2	S77457	sensory transduction	2.25e+02	335	5	17.2	558	2	S29125	dimethylalliline monoo	2.25e+02
263	5	17.2	440	2	H69989	lipoprotein homolog y	2.25e+02	336	5	17.2	560	2	F70688	hypothetical protein	2.25e+02
264	5	17.2	440	2	C55521	viRS protein - Clostr	2.25e+02	337	5	17.2	561	2	S73547	hypothetical protein	2.25e+02
265	5	17.2	446	2	A40896	Ca2+/calmodulin-depen	2.25e+02	338	5	17.2	568	2	I58106	gene DMR-N9 protein -	2.25e+02
266	5	17.2	447	2	S59742	SLE1 protein - yeast	2.25e+02	339	5	17.2	568	2	S19031	Poly(A) polymerase -	2.25e+02
267	5	17.2	448	2	T01814	hypothetical protein	2.25e+02	340	5	17.2	572	2	S49885	TOAD-64 protein - bla	2.25e+02
268	5	17.2	449	2	A24993	cellulase (EC 3.2.1.4	2.25e+02	341	5	17.2	572	2	S58889	collapsin response me	2.25e+02
269	5	17.2	449	2	A35762	mannosyltransferase (	2.25e+02	342	5	17.2	572	2	JC5317	dihydropyrimidinase r	2.25e+02
270	5	17.2	451	1	FOFFGY	retrovirus-related ga	2.25e+02	343	5	17.2	572	2	JC5316	dihydropyrimidinase r	2.25e+02
271	5	17.2	452	2	I49542	bone morphogenetic pr	2.25e+02	344	5	17.2	575	2	H69450	prolyl-tRNA synthetas	2.25e+02
272	5	17.2	453	2	A48348	genome polyprotein -	2.25e+02	345	5	17.2	575	2	S50962	hypothetical protein	2.25e+02
273	5	17.2	454	1	BMH05	bone morphogenetic pr	2.25e+02	346	5	17.2	578	2	A64696	hypothetical protein	2.25e+02
274	5	17.2	455	2	A40163	glycylpeptide N-tetra	2.25e+02	347	5	17.2	578	2	D64993	yfaA protein - Escher	2.25e+02
275	5	17.2	455	2	S56695	1-aminocyclopropane-1	2.25e+02	348	5	17.2	588	2	A26158	decapentaplegic prote	2.25e+02
276	5	17.2	459	2	S33000	hypothetical protein	2.25e+02	349	5	17.2	589	1	ORCHE	estrogen receptor - c	2.25e+02
277	5	17.2	459	2	S64241	hypothetical protein	2.25e+02	350	5	17.2	590	2	T01237	homeotic receptor, ovu	2.25e+02
278	5	17.2	460	2	S35525	telomere-binding prot	2.25e+02	351	5	17.2	595	2	I47140	estradiol receptor -	2.25e+02
279	5	17.2	464	2	A34625	protein-tyrosine kina	2.25e+02	352	5	17.2	597	2	D67111	glutamine--fructose-6	2.25e+02
280	5	17.2	465	2	JC5184	glutamate synthase (G	2.25e+02	353	5	17.2	599	1	QRMSE	estrogen receptor - m	2.25e+02
281	5	17.2	465	2	S70879	GDP-mannose pyrophosp	2.25e+02	354	5	17.2	600	1	ORRTE	estrogen receptor - r	2.25e+02
282	5	17.2	467	1	A30042	transcription factor	2.25e+02	355	5	17.2	604	2	S69793	glutamine--fructose-6	2.25e+02
283	5	17.2	468	2	H71932	myosin VIIA - human (	2.25e+02	356	5	17.2	605	2	S48940	hypothetical protein	2.25e+02
284	5	17.2	470	2	H70795	hypothetical protein	2.25e+02	357	5	17.2	611	1	S06047	endo-1,4-beta-xylans	2.25e+02
285	5	17.2	470	2	G70632	probable NADH DEH/DRO	2.25e+02	358	5	17.2	614	2	A32608	thyroid hormone recep	2.25e+02
286	5	17.2	471	2	S51511	capsanthin/capsorubin	2.25e+02	359	5	17.2	617	2	A34174	secretogranin II prec	2.25e+02
287	5	17.2	472	2	S31459	Ig gamma-1 chain - sh	2.25e+02	360	5	17.2	617	2	S27389	secretogranin II - mo	2.25e+02
288	5	17.2	477	2	S71400	estrogen receptor bet	2.25e+02	361	5	17.2	619	2	S02180	secretogranin II - ra	2.25e+02
289	5	17.2	478	2	C69743	conserved hypothetica	2.25e+02	362	5	17.2	623	2	S59234	glutamate--cysteine l	2.25e+02
290	5	17.2	481	2	A46602	glutamate synthase (N	2.25e+02	363	5	17.2	624	2	JW0087	sulfate adenylyltrans	2.25e+02
291	5	17.2	482	2	G71345	conserved hypothetica	2.25e+02	364	5	17.2	626	2	S36364	glucan 1,4-alpha-gluc	2.25e+02
292	5	17.2	485	2	A35516	1-aminocyclopropane-1	2.25e+02	365	5	17.2	631	2	S60346	myosin VII - mouse (f	2.25e+02
293	5	17.2	485	2	S19677	1-aminocyclopropane-1	2.25e+02	366	5	17.2	633	2	T02686	hypothetical protein	2.25e+02
294	5	17.2	489	2	S58725	dual specificity phos	2.25e+02	367	5	17.2	634	2	T00388	hypothetical protein	2.25e+02
295	5	17.2	490	2	H70064	permease homolog ywoE	2.25e+02	368	5	17.2	636	2	S63131	probable membrane pro	2.25e+02
296	5	17.2	491	2	S48827	1-aminocyclopropane-1	2.25e+02	369	5	17.2	637	2	T01961	hypothetical protein	2.25e+02
297	5	17.2	493	2	B70417	conserved hypothetica	2.25e+02	370	5	17.2	644	2	T07094	probable alpha-isopro	2.25e+02
298	5	17.2	493	2	S63060	hypothetical protein	2.25e+02	371	5	17.2	651	2	A39372	potassium channel pr	2.25e+02
299	5	17.2	494	2	S15523	variant surface glyco	2.25e+02	372	5	17.2	652	2	S71753	repellent protein l p	2.25e+02
300	5	17.2	495	2	A44489	GT box-binding protei	2.25e+02	373	5	17.2	658	1	S73805	DNA ligase (NAD+) (EC	2.25e+02
301	5	17.2	498	2	S78184	NADH dehydrogenase (u	2.25e+02	374	5	17.2	658	2	S60170	protein kinase Pak1 -	2.25e+02
302	5	17.2	498	2	S71511	capsanthin/capsorubin	2.25e+02	375	5	17.2	659	2	S46788	PES4 protein homolog	2.25e+02
303	5	17.2	501	1	PLWLEP	Li protein - European	2.25e+02	376	5	17.2	665	2	T00015	unc-14 protein - Caen	2.25e+02
304	5	17.2	503	2	S64787	pre-mRNA splicing pro	2.25e+02	377	5	17.2	667	2	T07019	probable carbamoyl-ph	2.25e+02
305	5	17.2	504	2	B40829	activin receptor isof	2.25e+02	378	5	17.2	671	1	ORBOG	protein kinase (EC 2.	2.25e+02
306	5	17.2	509	2	S58890	collapsin response me	2.25e+02	379	5	17.2	673	2	T00328	hypothetical protein	2.25e+02
307	5	17.2	510	2	JX0254	embryonic nuclear pro	2.25e+02	380	5	17.2	676	2	A45984	sperm-binding glycopr	2.25e+02
308	5	17.2	510	1	A42635	activin receptor strX8	2.25e+02	381	5	17.2	686	2	S66693	hypothetical protein	2.25e+02
309	5	17.2	512	2	I37134	activin type II recep	2.25e+02	382	5	17.2	686	1	S05702	protein kinase (EC 2.	2.25e+02
310	5	17.2	512	2	D40829	activin receptor isof	2.25e+02	383	5	17.2	690	2	C71982	probable outer membra	2.25e+02
311	5	17.2	512	1	RDECNB	nitrate reductase (EC	2.25e+02	384	5	17.2	696	2	A41344	lutropin-choriogonada	2.25e+02
312	5	17.2	513	2	JQ1484	activin receptor prec	2.25e+02	385	5	17.2	698	2	S52674	general sporulation p	2.25e+02
313	5	17.2	513	1	PLWLDP	Li protein - deer pap	2.25e+02	386	5	17.2	699	2	I51893	smooth muscle myosin	2.25e+02
314	5	17.2	518	2	A55066	tyrosine decarboxylas	2.25e+02	387	5	17.2	699	1	QRHUUT	lutropin-choriogonada	2.25e+02
315	5	17.2	520	2	S57337	trichodiene oxygenase	2.25e+02	388	5	17.2	700	2	S57415	Hsp83 protein - Leish	2.25e+02

389	5	17.2	701	2	A44943	2.25e+02	462	5	17.2	917	2	S35228	nitrate reductase (NA	2.25e+02
390	5	17.2	701	2	S64599	2.25e+02	463	5	17.2	918	2	A41667	nitrate reductase (NA	2.25e+02
391	5	17.2	702	2	A48754	2.25e+02	464	5	17.2	920	2	S52301	nitrate reductase (NA	2.25e+02
392	5	17.2	705	2	A41322	2.25e+02	465	5	17.2	923	2	S44664	ZK370.3 protein - Cae	2.25e+02
393	5	17.2	706	2	I64833	2.25e+02	466	5	17.2	926	1	RDSPNH	nitrate reductase (NA	2.25e+02
394	5	17.2	706	2	S53035	2.25e+02	467	5	17.2	927	2	J00948	A5 antigen precursor	2.25e+02
395	5	17.2	708	2	A56163	2.25e+02	468	5	17.2	940	2	J00291	FB19 protein - human	2.25e+02
396	5	17.2	717	1	HHPF83	2.25e+02	469	5	17.2	946	2	S66367	H+-transporting ATPase	2.25e+02
397	5	17.2	718	2	S71211	2.25e+02	470	5	17.2	952	1	QYMG	phosphoenolpyruvate c	2.25e+02
398	5	17.2	719	2	A42608	2.25e+02	471	5	17.2	960	2	S22507	phosphoenolpyruvate c	2.25e+02
399	5	17.2	720	2	D70356	2.25e+02	472	5	17.2	1011	1	GNVYC1	genome polyprotein -	2.25e+02
400	5	17.2	722	2	B61231	2.25e+02	473	5	17.2	1016	2	J00498	genome polyprotein -	2.25e+02
401	5	17.2	727	2	A38206	2.25e+02	474	5	17.2	1018	2	A32192	fibronectin-binding p	2.25e+02
402	5	17.2	732	2	S47688	2.25e+02	475	5	17.2	1022	1	S00503	Na+/K+-exchanging ATP	2.25e+02
403	5	17.2	733	2	J01892	2.25e+02	476	5	17.2	1030	2	S11034	gene p1 protein - fru	2.25e+02
404	5	17.2	733	2	J01891	2.25e+02	477	5	17.2	1039	2	S76747	hypothetical protein	2.25e+02
405	5	17.2	740	2	JC5712	2.25e+02	478	5	17.2	1039	2	S62509	hypothetical protein	2.25e+02
406	5	17.2	746	2	T01536	2.25e+02	479	5	17.2	1047	2	D71302	probable exonuclease	2.25e+02
407	5	17.2	757	2	J50198	2.25e+02	480	5	17.2	1064	2	T01960	hypothetical protein	2.25e+02
408	5	17.2	758	2	S45825	2.25e+02	481	5	17.2	1090	2	T00533	probable trRNA-splicin	2.25e+02
409	5	17.2	762	1	NNNC2	2.25e+02	482	5	17.2	1091	2	S33596	protein-tyrosine kina	2.25e+02
410	5	17.2	766	2	G71437	2.25e+02	483	5	17.2	1120	2	H64776	probable membrane pro	2.25e+02
411	5	17.2	768	2	S71477	2.25e+02	484	5	17.2	1132	2	S60433	probable membrane pro	2.25e+02
412	5	17.2	776	1	RGECAR	2.25e+02	485	5	17.2	1134	2	G92699	molybdopter-in-binding	2.25e+02
413	5	17.2	783	2	A48998	2.25e+02	486	5	17.2	1146	2	S46837	hypothetical protein	2.25e+02
414	5	17.2	783	2	T00354	2.25e+02	487	5	17.2	1154	2	S39536	parasporal crystal pr	2.25e+02
415	5	17.2	790	2	S27458	2.25e+02	488	5	17.2	1163	2	A56097	arylphorin-binding pr	2.25e+02
416	5	17.2	796	2	H71865	2.25e+02	489	5	17.2	1169	1	S38181	flocculation protein	2.25e+02
417	5	17.2	796	2	G64570	2.25e+02	490	5	17.2	1169	1	S64859	DNA repair protein RA	2.25e+02
418	5	17.2	797	2	D71621	2.25e+02	491	5	17.2	1177	2	A36984	protein-tyrosine kina	2.25e+02
419	5	17.2	801	1	MXRR5	2.25e+02	492	5	17.2	1188	2	A71621	protein with 5'-3' ex	2.25e+02
420	5	17.2	806	2	S76650	2.25e+02	493	5	17.2	1220	2	S64916	probable membrane pro	2.25e+02
421	5	17.2	810	2	I46260	2.25e+02	494	5	17.2	1244	2	S76102	hypothetical protein	2.25e+02
422	5	17.2	823	2	S53647	2.25e+02	495	5	17.2	1246	2	J00406	hypothetical protein	2.25e+02
423	5	17.2	824	2	I52835	2.25e+02	496	5	17.2	1260	2	S80896	agglutinin-like prote	2.25e+02
424	5	17.2	830	2	A34347	2.25e+02	497	5	17.2	1297	2	S39791	neurotoxin - Clostrid	2.25e+02
425	5	17.2	835	2	JC6140	2.25e+02	498	5	17.2	1336	2	S41794	SEC3 protein - yeast	2.25e+02
426	5	17.2	842	2	A41778	2.25e+02	499	5	17.2	1371	2	S77521	sensory transduction	2.25e+02
427	5	17.2	851	2	S12159	2.25e+02	500	5	17.2	1374	1	GNVYC	genome polyprotein -	2.25e+02
428	5	17.2	852	1	VCLJGG	2.25e+02	501	5	17.2	1375	2	JC5148	hepatocyte growth fac	2.25e+02
429	5	17.2	854	2	S61164	2.25e+02	502	5	17.2	1379	2	S01254	hepatocyte growth fac	2.25e+02
430	5	17.2	855	2	A48168	2.25e+02	503	5	17.2	1390	1	TVHUME	hepatocyte growth fac	2.25e+02
431	5	17.2	858	1	VCLJG2	2.25e+02	504	5	17.2	1444	1	A43377	RNA-directed RNA poly	2.25e+02
432	5	17.2	859	1	VCLJST	2.25e+02	505	5	17.2	1569	2	G02757	clathrin heavy chain	2.25e+02
433	5	17.2	859	1	VCLJST	2.25e+02	506	5	17.2	1592	2	A38175	glucosyltransferase p	2.25e+02
434	5	17.2	860	2	G70623	2.25e+02	507	5	17.2	1597	2	S65053	genome polyprotein -	2.25e+02
435	5	17.2	863	2	S38140	2.25e+02	508	5	17.2	1675	1	LRRTH	clathrin heavy chain	2.25e+02
436	5	17.2	863	2	A48902	2.25e+02	509	5	17.2	1790	2	S67593	transport protein USO	2.25e+02
437	5	17.2	863	2	A53034	2.25e+02	510	5	17.2	1805	2	A34736	nestin - rat	2.25e+02
438	5	17.2	864	2	S63229	2.25e+02	511	5	17.2	1827	2	B70984	probable ppsd protein	2.25e+02
439	5	17.2	864	2	S07075	2.25e+02	512	5	17.2	1848	2	A44140	cellulose-binding pro	2.25e+02
440	5	17.2	864	2	A48266	2.25e+02	513	5	17.2	1850	1	VUCH2	vitellogenin II precu	2.25e+02
441	5	17.2	869	2	S53098	2.25e+02	514	5	17.2	1852	2	I50441	vitellogenin - chicke	2.25e+02
442	5	17.2	875	2	I59350	2.25e+02	515	5	17.2	1938	2	JC5421	smooth muscle myosin	2.25e+02
443	5	17.2	876	2	I52907	2.25e+02	516	5	17.2	1959	1	A33977	myosin heavy chain, n	2.25e+02
444	5	17.2	876	2	S66288	2.25e+02	517	5	17.2	1961	1	A61231	myosin heavy chain, n	2.25e+02
445	5	17.2	877	2	A49197	2.25e+02	518	5	17.2	1972	1	A41604	myosin heavy chain, s	2.25e+02
446	5	17.2	879	2	A55881	2.25e+02	519	5	17.2	1972	2	JC5420	smooth muscle myosin	2.25e+02
447	5	17.2	880	2	S72503	2.25e+02	520	5	17.2	1979	1	S03166	smooth muscle myosin	2.25e+02
448	5	17.2	880	2	S66307	2.25e+02	521	5	17.2	1983	2	T00385	KIAA0624 protein - hu	2.25e+02
449	5	17.2	881	2	S25445	2.25e+02	522	5	17.2	1992	2	A72797	myosin heavy chain fo	2.25e+02
450	5	17.2	884	2	S66308	2.25e+02	523	5	17.2	1999	1	S21801	myosin heavy chain, n	2.25e+02
451	5	17.2	887	1	DESCPV	2.25e+02	524	5	17.2	2007	1	B43402	myosin heavy chain-B,	2.25e+02
452	5	17.2	893	2	S63378	2.25e+02	525	5	17.2	2017	1	A36014	myosin heavy chain, n	2.25e+02
453	5	17.2	898	1	RDBJNH	2.25e+02	526	5	17.2	2026	1	ORBY	adenylate cyclase [EC	2.25e+02
454	5	17.2	900	2	S25322	2.25e+02	527	5	17.2	2057	2	S61477	myosin II heavy chain	2.25e+02
455	5	17.2	900	2	S47029	2.25e+02	528	5	17.2	2092	2	S30026	genome polyprotein -	2.25e+02
456	5	17.2	904	1	RDNTNT	2.25e+02	529	5	17.2	2116	1	ZLVNSY	genome polyprotein -	2.25e+02
457	5	17.2	904	1	RDNTNS	2.25e+02	530	5	17.2	2139	2	S46404	vitellogenin - yellow	2.25e+02
458	5	17.2	907	2	A45560	2.25e+02	531	5	17.2	2149	2	S18676	genome polyprotein -	2.25e+02
459	5	17.2	909	2	JN0665	2.25e+02	532	5	17.2	2150	1	GNVYH2	genome polyprotein -	2.25e+02
460	5	17.2	911	1	RDTONH	2.25e+02	533	5	17.2	2182	1	GNVYB1	genome polyprotein -	2.25e+02
461	5	17.2	913	2	D71463	2.25e+02	534	5	17.2	2183	1	GNVYB4	genome polyprotein -	2.25e+02

535	5	17.2	2185	1	JQ2021	genome polypeptide -	2.25e+02	608	4	13.8	127	1	VBRB	transferrin - rabbi	4.36e+03
536	5	17.2	2185	1	GNYSH	genome polypeptide -	2.25e+02	609	4	13.8	129	2	S03534	Ig heavy chain precu	4.36e+03
537	5	17.2	2185	1	GNYSB	genome polypeptide -	2.25e+02	610	4	13.8	133	2	S73222	ribosomal protein S8	4.36e+03
538	5	17.2	2185	1	GNYSV	genome polypeptide -	2.25e+02	611	4	13.8	135	2	G71547	probable protein-disu	4.36e+03
539	5	17.2	2185	1	GNYSV	genome polypeptide -	2.25e+02	612	4	13.8	135	2	E69960	conserved hypothetica	4.36e+03
540	5	17.2	2193	2	S44194	genome polypeptide -	2.25e+02	613	4	13.8	136	2	B45893	T-cell receptor alpha	4.36e+03
541	5	17.2	2193	2	S44251	polyprotein - echovir	2.25e+02	614	4	13.8	140	2	S54227	Ig mu heavy chain v r	4.36e+03
542	5	17.2	2193	2	S52919	polyprotein - echovir	2.25e+02	615	4	13.8	140	2	S54239	Ig mu heavy chain v r	4.36e+03
543	5	17.2	2194	1	GNYS7	genome polypeptide -	2.25e+02	616	4	13.8	140	2	S54245	Ig mu heavy chain v r	4.36e+03
544	5	17.2	2201	1	GNYS9	genome polypeptide -	2.25e+02	617	4	13.8	140	2	I37782	Ig variable region (v	4.36e+03
545	5	17.2	2231	1	S53416	SEN1 protein - yeast	2.25e+02	618	4	13.8	141	2	S77320	hypothetical protein	4.36e+03
546	5	17.2	2284	1	GNYSV	genome polypeptide -	2.25e+02	619	4	13.8	141	2	S54238	Ig mu heavy chain v r	4.36e+03
547	5	17.2	2325	2	A61208	chondroitin sulfate p	2.25e+02	620	4	13.8	142	2	S03995	hemoglobin IV alpha c	4.36e+03
548	5	17.2	2332	1	GNYSF	genome polypeptide -	2.25e+02	621	4	13.8	142	2	A60514	hemoglobin 2 alpha ch	4.36e+03
549	5	17.2	2332	1	GNYSF	genome polypeptide -	2.25e+02	622	4	13.8	144	2	S54244	Ig mu heavy chain v r	4.36e+03
550	5	17.2	2333	1	GNYS2F	genome polypeptide -	2.25e+02	623	4	13.8	145	2	B58794	hemoglobin beta chain	4.36e+03
551	5	17.2	2336	2	S37077	genome polypeptide -	2.25e+02	624	4	13.8	145	2	S77259	large conductance mec	4.36e+03
552	5	17.2	2339	2	A45597	DNA-directed RNA poly	2.25e+02	625	4	13.8	146	1	HBKKS	hemoglobin beta chain	4.36e+03
553	5	17.2	2467	2	D71437	probable resistance g	2.25e+02	626	4	13.8	146	1	HBKKN	hemoglobin beta chain	4.36e+03
554	5	17.2	2493	2	S45734	probable membrane pro	2.25e+02	627	4	13.8	146	1	HGMQJ	hemoglobin gamma cha	4.36e+03
555	5	17.2	2515	2	A41519	posterior-group prote	2.25e+02	628	4	13.8	146	1	HBCZP	hemoglobin beta chain	4.36e+03
556	5	17.2	2833	2	A43360	inositol 1,4,5-trisph	2.25e+02	629	4	13.8	146	2	B25729	hemoglobin beta chain	4.36e+03
557	5	17.2	2875	1	RRVUTW	genome polypeptide -	2.25e+02	630	4	13.8	146	1	HBMQP	hemoglobin beta chain	4.36e+03
558	5	17.2	2960	2	A45259	desmoyokin - human (f	2.25e+02	631	4	13.8	146	2	S30434	Ig heavy chain v-D-J	4.36e+03
559	5	17.2	3066	1	JQ1662	genome polypeptide -	2.25e+02	632	4	13.8	146	2	B23391	hemoglobin beta chain	4.36e+03
560	5	17.2	3066	1	JQ1661	genome polypeptide -	2.25e+02	633	4	13.8	147	2	I84426	hemoglobin epsilon ch	4.36e+03
561	5	17.2	3591	2	S10236	filamentous hemagglut	2.25e+02	634	4	13.8	147	2	I37054	hemoglobin epsilon ch	4.36e+03
562	5	17.2	3591	2	S21010	filamentous hemagglut	2.25e+02	635	4	13.8	147	1	HGMQR2	hemoglobin gamma-2 ch	4.36e+03
563	5	17.2	3795	2	T00831	hypothetical protein	2.25e+02	636	4	13.8	147	1	HBCZ	hemoglobin beta chain	4.36e+03
564	5	17.2	4199	2	S76412	hypothetical protein	2.25e+02	637	4	13.8	147	2	B42829	hemoglobin gamma cha	4.36e+03
565	5	17.2	4351	2	T00252	MEGF1 protein - rat	2.25e+02	638	4	13.8	150	2	B64020	hypothetical protein	4.36e+03
566	5	17.2	4344	1	S03392	alpha-2-macroglobulin	2.25e+02	639	4	13.8	150	2	S09872	hypothetical protein	4.36e+03
567	5	17.2	4545	1	S25111	alpha-2-macroglobulin	2.25e+02	640	4	13.8	152	2	S41727	unc-60 protein - Caen	4.36e+03
568	5	17.2	26926	1	I38344	titin, cardiac muscle	2.25e+02	641	4	13.8	153	1	A42156	translation initiatio	4.36e+03
569	4	13.8	15	2	C48401	ribosomal protein L1	4.36e+03	642	4	13.8	155	2	S17912	NADH dehydrogenase (u	4.36e+03
570	4	13.8	46	2	I39904	GTP cyclohydrolase I	4.36e+03	643	4	13.8	155	2	S56376	sugES protein - Esche	4.36e+03
571	4	13.8	48	2	A38882	rab protein cdc42 - m	4.36e+03	644	4	13.8	162	2	S31073	R-phycocyanin II alph	4.36e+03
572	4	13.8	69	1	Q0V218	hypothetical protein	4.36e+03	645	4	13.8	162	2	S74345	hypothetical protein	4.36e+03
573	4	13.8	77	2	S35869	envelope protein - hu	4.36e+03	646	4	13.8	166	2	B57378	cyclin-dependent kina	4.36e+03
574	4	13.8	77	2	S35815	envelope protein - hu	4.36e+03	647	4	13.8	167	1	JC1102	endothelin 3 precurs	4.36e+03
575	4	13.8	77	2	S35831	envelope protein - hu	4.36e+03	648	4	13.8	169	1	FIDOA	translation initiatio	4.36e+03
576	4	13.8	83	2	JN0400	hypothetical 8.8K pro	4.36e+03	649	4	13.8	176	2	A64497	major allergen I - As	4.36e+03
577	4	13.8	83	2	JQ1962	hypothetical 9.3K pro	4.36e+03	650	4	13.8	182	1	RKJ4M	ribulose-bisphosphate	4.36e+03
578	4	13.8	97	2	S04061	matrix protein M2 - i	4.36e+03	651	4	13.8	186	2	A29802	napin precursor (gna)	4.36e+03
579	4	13.8	97	2	S14517	matrix protein M2 - u	4.36e+03	652	4	13.8	187	2	S23595	embryonic fibroblast	4.36e+03
580	4	13.8	100	2	S12433	Ig heavy chain v regi	4.36e+03	653	4	13.8	188	2	I64061	translation elongatio	4.36e+03
581	4	13.8	100	1	R3SP14	ribosomal protein S14	4.36e+03	654	4	13.8	190	2	A70100	hypothetical protein	4.36e+03
582	4	13.8	101	2	S56419	hypothetical 10.9K pr	4.36e+03	655	4	13.8	191	2	S57563	CD42 GTP-binding pro	4.36e+03
583	4	13.8	102	2	S14488	Ig heavy chain v regi	4.36e+03	656	4	13.8	192	2	S42963	viral infectivity fac	4.36e+03
584	4	13.8	103	2	A25875	histone H4 - Tetrahym	4.36e+03	657	4	13.8	194	2	S75738	aminoacyl-tRNA hydrol	4.36e+03
585	4	13.8	104	2	S14184	histone H4 (clone H4K	4.36e+03	658	4	13.8	194	2	A25327	adenylate kinase (EC	4.36e+03
586	4	13.8	105	1	WRBP68	early protein gp12 -	4.36e+03	659	4	13.8	199	2	S25967	NADH dehydrogenase (u	4.36e+03
587	4	13.8	107	2	A94442	Ig heavy chain v regi	4.36e+03	660	4	13.8	200	1	MORT3V	myosin alkali light c	4.36e+03
588	4	13.8	111	2	S62830	ribosomal protein L24	4.36e+03	661	4	13.8	202	1	DSPMC2	superoxide dismutase	4.36e+03
589	4	13.8	113	2	S28465	Ig heavy chain v regi	4.36e+03	662	4	13.8	203	1	WMPTB	gene 10 protein - pha	4.36e+03
590	4	13.8	114	2	S43714	ferredoxin--thioredox	4.36e+03	663	4	13.8	204	2	C40899	hypothetical protein	4.36e+03
591	4	13.8	114	2	S22554	Ig heavy chain v regi	4.36e+03	664	4	13.8	206	2	S55604	hypothetical protein	4.36e+03
592	4	13.8	115	1	R6RT30	ribosomal protein L30	4.36e+03	665	4	13.8	207	1	LNCHL	hepatic lectin - chic	4.36e+03
593	4	13.8	115	2	B32071	T-cell receptor delta	4.36e+03	666	4	13.8	208	2	S14700	hypothetical protein	4.36e+03
594	4	13.8	115	2	S34608	ribosomal protein L30	4.36e+03	667	4	13.8	211	2	D69765	lactam utilization pr	4.36e+03
595	4	13.8	116	1	R5HSH9	ribosomal protein H12	4.36e+03	668	4	13.8	216	2	JQ1841	10R protein - variola	4.36e+03
596	4	13.8	118	2	S24702	Ig heavy chain V6 reg	4.36e+03	669	4	13.8	218	2	F69633	glutamine ABC transpo	4.36e+03
597	4	13.8	119	2	S24662	Ig heavy chain v regi	4.36e+03	670	4	13.8	222	2	C65001	hypothetical protein	4.36e+03
598	4	13.8	121	1	YQSECF	fimbrial protein prec	4.36e+03	671	4	13.8	223	1	YLHUP	serum amyloid P-compo	4.36e+03
599	4	13.8	122	1	Z12M49	22K zein (clone B49)	4.36e+03	672	4	13.8	224	2	A64055	GTP cyclohydrolase II	4.36e+03
600	4	13.8	123	2	S24691	Ig heavy chain V6 reg	4.36e+03	673	4	13.8	225	1	WMNV29	DA26 protein - Auto9	4.36e+03
601	4	13.8	123	2	S24694	Ig heavy chain V6 reg	4.36e+03	674	4	13.8	225	2	S39712	uracil-DNA glycosylas	4.36e+03
602	4	13.8	124	2	S24653	Ig heavy chain V6 reg	4.36e+03	675	4	13.8	226	2	JE0151	myonase (EC 3.4.-.-)	4.36e+03
603	4	13.8	125	2	S24703	Ig heavy chain V6 reg	4.36e+03	676	4	13.8	226	2	S46353	nef protein - simian	4.36e+03
604	4	13.8	125	2	S24700	Ig heavy chain V6 reg	4.36e+03	677	4	13.8	227	1	S68150	eosinophil major basi	4.36e+03
605	4	13.8	126	2	S24704	Ig heavy chain V6 reg	4.36e+03	678	4	13.8	228	2	S26995	cellulose 1,4-beta-ce	4.36e+03
606	4	13.8	126	2	JC2034	pancreatic ribonuclea	4.36e+03	679	4	13.8	230	2	A64387	fibrillar-like pre-	4.36e+03
607	4	13.8	127	2	S54268	fatty acid-binding pr	4.36e+03	680	4	13.8	236	2	A24637	T-cell surface glycop	4.36e+03

681	4	13.8	243	2	JC4585	GTP cyclohydrolase I	4.36e+03	754	4	13.8	376	2	S47986	actin-related protein	4.36e+03
682	4	13.8	247	2	S10174	H+-transporting ATP s	4.36e+03	755	4	13.8	381	2	S51375	microtubule-associate	4.36e+03
683	4	13.8	247	1	LNR26	H+-transporting ATP s	4.36e+03	756	4	13.8	384	2	S20303	substance K receptor	4.36e+03
684	4	13.8	248	2	A6322	3-deoxy-manno-otulos	4.36e+03	757	4	13.8	385	2	E59719	succinate-CoA ligase	4.36e+03
685	4	13.8	250	2	S1384	trypsin (EC 3.4.21.4)	4.36e+03	758	4	13.8	390	2	D71283	conserved hypothetica	4.36e+03
686	4	13.8	251	2	A36845	M4R protein - variola	4.36e+03	759	4	13.8	391	1	J01626	attachment protein G	4.36e+03
687	4	13.8	251	2	B42513	L4R protein - vaccini	4.36e+03	760	4	13.8	393	1	XXALAE	acetyl-CoA C-acetyltr	4.36e+03
688	4	13.8	252	2	S25965	H+-transporting ATP s	4.36e+03	761	4	13.8	395	2	F64534	tail sheath protein -	4.36e+03
689	4	13.8	261	1	WHUET	proliferating cell nu	4.36e+03	762	4	13.8	396	2	S26391	hypothetical protein	4.36e+03
690	4	13.8	264	1	WBVL1	30K protein - tomato	4.36e+03	763	4	13.8	397	1	XURT	acetyl-CoA C-acyltran	4.36e+03
691	4	13.8	273	1	RDECPD	dihydrodipicolinate r	4.36e+03	764	4	13.8	400	2	I47099	renin (EC 3.4.23.11)	4.36e+03
692	4	13.8	275	2	C44054	orf3 protein - Junoni	4.36e+03	765	4	13.8	401	2	C71310	conserved hypothetica	4.36e+03
693	4	13.8	278	2	A41700	D9 protein - rabbit f	4.36e+03	766	4	13.8	402	1	F48552	protein kinase (EC 2.	4.36e+03
694	4	13.8	280	2	B45537	viral coat protein -	4.36e+03	767	4	13.8	405	1	CBQFR	ubiquinol--cytochrome	4.36e+03
695	4	13.8	280	2	S76022	aspartate-semialdehyd	4.36e+03	768	4	13.8	405	2	A64388	3-hydroxy-3-methylglu	4.36e+03
696	4	13.8	281	2	B44227	hypothetical protein	4.36e+03	769	4	13.8	407	2	S47680	CDC2/CDC13 suppressor	4.36e+03
697	4	13.8	282	2	A45537	viral coat protein -	4.36e+03	770	4	13.8	413	2	S21560	aspartate transaminas	4.36e+03
698	4	13.8	283	2	C70478	proteinase IV - Aquif	4.36e+03	771	4	13.8	415	2	S15709	hypothetical protein	4.36e+03
699	4	13.8	284	2	S74168	tropomyosin - America	4.36e+03	772	4	13.8	416	2	JU0475	phosphoglycerate kina	4.36e+03
700	4	13.8	285	2	S62724	ribulose-phosphate 3-	4.36e+03	773	4	13.8	416	2	A45105	farnesyl-diphosphate	4.36e+03
701	4	13.8	288	2	J80659	homeotic protein Hox-	4.36e+03	774	4	13.8	419	2	S71343	calreticulin precursor	4.36e+03
702	4	13.8	291	2	D84043	citrate (pro-3S)-lyas	4.36e+03	775	4	13.8	423	1	EHMS5	Ig epsilon chain C re	4.36e+03
703	4	13.8	293	2	A40131	T-cell receptor delta	4.36e+03	776	4	13.8	423	2	H70310	ammonium transporter	4.36e+03
704	4	13.8	296	2	H65030	hypothetical protein	4.36e+03	777	4	13.8	424	1	XUH0AB	acetyl-CoA C-acyltran	4.36e+03
705	4	13.8	297	2	S42144	ribosomal protein L5.	4.36e+03	778	4	13.8	426	1	A48829	transcription factor	4.36e+03
706	4	13.8	297	1	MNVNRV	nonstructural protein	4.36e+03	779	4	13.8	427	2	S75210	glycine hydroxymethyl	4.36e+03
707	4	13.8	297	2	S63700	ubiquinol--cytochrome	4.36e+03	780	4	13.8	427	2	A71612	eukaryotic peptide ch	4.36e+03
708	4	13.8	299	1	NAIF	nitrogenase (EC 1.18.	4.36e+03	781	4	13.8	428	2	JQ1864	hypothetical 47.0K pr	4.36e+03
709	4	13.8	301	2	A32814	flagellar filament su	4.36e+03	782	4	13.8	428	2	S56679	mitosis-specific cycl	4.36e+03
710	4	13.8	302	2	A47126	alsSD operon activato	4.36e+03	783	4	13.8	429	1	A4ECCG	phosphoribosylamine--	4.36e+03
711	4	13.8	303	2	C33374	hypothetical protein	4.36e+03	784	4	13.8	430	1	A46216	transcription factor	4.36e+03
712	4	13.8	308	2	B70103	methylenetetrahydrof	4.36e+03	785	4	13.8	431	1	JC2002	transcription factor	4.36e+03
713	4	13.8	311	2	S35089	modulation formatio-	4.36e+03	786	4	13.8	432	2	H64152	hypothetical protein	4.36e+03
714	4	13.8	312	2	D89376	conserved hypothetica	4.36e+03	787	4	13.8	437	1	B29336	ubiquinol--cytochrome	4.36e+03
715	4	13.8	317	2	E59763	probable ferrichrome	4.36e+03	788	4	13.8	438	2	JC1179	transforming protein	4.36e+03
716	4	13.8	318	2	S41290	adenosine receptor A3	4.36e+03	789	4	13.8	439	2	S16576	S-locus-specific glyc	4.36e+03
717	4	13.8	323	2	S42426	aspartate-semialdehyd	4.36e+03	790	4	13.8	440	2	I48291	transcription factor	4.36e+03
718	4	13.8	325	2	A33851	alcohol dehydrogenase	4.36e+03	791	4	13.8	440	1	TVHUT	transforming protein	4.36e+03
719	4	13.8	326	2	P80017	Ig gamma-1 chain C re	4.36e+03	792	4	13.8	440	2	H64658	adenylosuccinate lyas	4.36e+03
720	4	13.8	326	2	J64124	pregnancy-specific G1	4.36e+03	793	4	13.8	440	2	B71858	adenylosuccinate lyas	4.36e+03
721	4	13.8	326	2	S32585	structural protein VP	4.36e+03	794	4	13.8	440	1	TWBYM1	transcription factor	4.36e+03
722	4	13.8	327	2	S25421	nucleolar protein NOP	4.36e+03	795	4	13.8	445	2	A44949	tubulin beta chain -	4.36e+03
723	4	13.8	327	2	S25421	malate dehydrogenase	4.36e+03	796	4	13.8	447	2	S44133	histidine kinase - La	4.36e+03
724	4	13.8	338	1	DETM5	malate dehydrogenase	4.36e+03	797	4	13.8	449	2	S28167	cytochrome P450 2E1 -	4.36e+03
725	4	13.8	338	1	DETM5	malate dehydrogenase	4.36e+03	798	4	13.8	450	2	S20868	tubulin beta-1 chain	4.36e+03
726	4	13.8	343	2	A52117	hypothetical protein	4.36e+03	799	4	13.8	451	2	S46048	hypothetical protein	4.36e+03
727	4	13.8	344	2	S09883	hypothetical protein	4.36e+03	800	4	13.8	454	2	I53224	p87 myc oncogene - hu	4.36e+03
728	4	13.8	351	2	S21345	retrovirus-related le	4.36e+03	801	4	13.8	455	1	TVCRGG	phosphoglycerate kina	4.36e+03
729	4	13.8	352	2	A70140	hypothetical protein	4.36e+03	802	4	13.8	457	2	A64711	chromosomal replicati	4.36e+03
730	4	13.8	353	1	J01946	core protein VP7 - Af	4.36e+03	803	4	13.8	459	2	A25928	cellulase (EC 3.2.1.4	4.36e+03
731	4	13.8	355	2	S76940	hypothetical protein	4.36e+03	804	4	13.8	462	1	COHU	hemopexin precursor -	4.36e+03
732	4	13.8	357	2	S22006	envelope protein gp12	4.36e+03	805	4	13.8	463	2	A48375	cellulase (EC 3.2.1.4	4.36e+03
733	4	13.8	358	2	B64427	carbamoyl-phosphate s	4.36e+03	806	4	13.8	465	2	S41084	triacylglycerol lipas	4.36e+03
734	4	13.8	358	2	S22000	envelope protein gp12	4.36e+03	807	4	13.8	466	2	A23685	interstitial collagen	4.36e+03
735	4	13.8	358	2	S70417	envelope protein gp12	4.36e+03	808	4	13.8	466	2	A64113	heat shock protein ht	4.36e+03
736	4	13.8	358	2	C70171	glycosyl transferase	4.36e+03	809	4	13.8	468	2	D64071	H+-transporting ATP s	4.36e+03
737	4	13.8	362	2	S73454	probable GTP-binding	4.36e+03	810	4	13.8	471	2	S68463	protein kinase ATPK19	4.36e+03
738	4	13.8	362	2	I61864	MHC HLA-B*41 chain -	4.36e+03	811	4	13.8	476	1	VYRTD	vitamin D-binding pro	4.36e+03
739	4	13.8	362	2	J73520	MHC class I histocomp	4.36e+03	812	4	13.8	478	2	E43599	toxin apxID secretion	4.36e+03
740	4	13.8	364	2	JC1188	GTP cyclohydrolase II	4.36e+03	813	4	13.8	479	2	S36841	EIL protein - variola	4.36e+03
741	4	13.8	366	1	A64107	DNA-directed DNA poly	4.36e+03	814	4	13.8	483	2	A33130	elav protein - fruit	4.36e+03
742	4	13.8	367	1	WOHU	alpha-2-HS-glycoprote	4.36e+03	815	4	13.8	484	2	A55626	monocarboxylate trans	4.36e+03
743	4	13.8	367	1	W2WL35	E2 protein - human pa	4.36e+03	816	4	13.8	485	1	A55731	GPI-anchor biosynthes	4.36e+03
744	4	13.8	368	1	VVP11L	coat protein VP1 - ly	4.36e+03	817	4	13.8	490	2	A2594	cytochrome P450 2C5,	4.36e+03
745	4	13.8	368	1	S64487	GTP-binding protein D	4.36e+03	818	4	13.8	493	2	S38621	actin bundling protei	4.36e+03
746	4	13.8	371	2	A71683	hypothetical protein	4.36e+03	819	4	13.8	496	2	C71920	aminopeptidase - HeII	4.36e+03
747	4	13.8	372	2	I39868	sac operon regulator	4.36e+03	820	4	13.8	497	2	C70454	transcription regulat	4.36e+03
748	4	13.8	374	2	S46264	microtubule-associate	4.36e+03	821	4	13.8	500	1	A44001	gag polyprotein - hum	4.36e+03
749	4	13.8	374	1	AS3142	alcohol dehydrogenase	4.36e+03	822	4	13.8	502	2	I40366	gag polyprotein - hum	4.36e+03
750	4	13.8	374	1	DEHUAS	alcohol dehydrogenase	4.36e+03	823	4	13.8	502	1	FOVW2	gag polyprotein - hum	4.36e+03
751	4	13.8	375	2	B71842	carbamoyl-phosphate s	4.36e+03	824	4	13.8	506	2	S76196	hypothetical protein	4.36e+03
752	4	13.8	375	2	A46174	RNA-binding protein r	4.36e+03	825	4	13.8	506	2	A40679	transcription enhance	4.36e+03
753	4	13.8	376	2	A56653	probable UTPase - hum	4.36e+03	826	4	13.8	510	2	A34160	laurate omega-hydroxy	4.36e+03

827	4	13.8	511	2	A39381	cytochrome P450 4 - C	4.36e+03	900	4	13.8	782	2	G64098	virulence-associated	4.36e+03
828	4	13.8	511	2	A71859	probable outer membra	4.36e+03	901	4	13.8	783	2	A55817	cyclin-dependent kina	4.36e+03
829	4	13.8	514	2	A44405	tyrosine decarboxylas	4.36e+03	902	4	13.8	788	2	S28302	ribonucleoside-diphos	4.36e+03
830	4	13.8	516	2	S43414	cytochrome P450 1A1 -	4.36e+03	903	4	13.8	800	2	A48991	heparin-binding growt	4.36e+03
831	4	13.8	519	2	T02263	cytochrome P450 DWARF	4.36e+03	904	4	13.8	804	1	YFBSB	phenylalanine--tRNA l	4.36e+03
832	4	13.8	520	2	S36495	L2 protein - human pa	4.36e+03	905	4	13.8	805	2	A69682	primosomal replicatio	4.36e+03
833	4	13.8	521	1	S35067	phosphoprotein phosph	4.36e+03	906	4	13.8	808	1	OPKEX	glucose dehydrogenase	4.36e+03
834	4	13.8	525	2	A70157	hypothetical protein	4.36e+03	907	4	13.8	810	2	A64742	hypothetical protein	4.36e+03
835	4	13.8	528	1	TVFVG9	protein-tyrosine kina	4.36e+03	908	4	13.8	810	1	A33380	interleukin-4 recepto	4.36e+03
836	4	13.8	531	2	B55066	tyrosine decarboxylas	4.36e+03	909	4	13.8	810	2	D64090	glycerol-3-phosphate	4.36e+03
837	4	13.8	537	2	I51592	p59(Xfyn) - Xiphophor	4.36e+03	910	4	13.8	838	2	B38656	vacuolar proton pump	4.36e+03
838	4	13.8	538	2	B68093	PRR2 delta - human	4.36e+03	911	4	13.8	838	2	A54163	vacuolar ATPase (EC 3	4.36e+03
839	4	13.8	542	2	A70220	phosphotransferase sy	4.36e+03	912	4	13.8	838	2	S23606	potassium channel pro	4.36e+03
840	4	13.8	544	1	B43627	chaperonin groEL - Br	4.36e+03	913	4	13.8	844	1	ZPECPB	penicillin-binding pr	4.36e+03
841	4	13.8	548	2	S61303	heat shock protein 63	4.36e+03	914	4	13.8	847	2	S13289	env protein - human i	4.36e+03
842	4	13.8	548	1	BVEGGL	chaperonin groEL - ES	4.36e+03	915	4	13.8	850	1	MMHUE4	erythrocyte membrane	4.36e+03
843	4	13.8	550	1	A26076	UDP-sugar hydrolase (	4.36e+03	916	4	13.8	852	2	D71896	aconitate hydratase -	4.36e+03
844	4	13.8	551	2	JC5225	dsRNA-activated prote	4.36e+03	917	4	13.8	852	2	I51259	tyrosine kinase C rec	4.36e+03
845	4	13.8	552	2	S43253	probable N3L protein	4.36e+03	918	4	13.8	853	2	C64617	aconitate hydratase (	4.36e+03
846	4	13.8	553	2	S71357	nucleation protein no	4.36e+03	919	4	13.8	854	2	S13288	env protein - human i	4.36e+03
847	4	13.8	556	2	I64068	phosphotransferase sy	4.36e+03	920	4	13.8	855	2	A45713	Env transmembrane pro	4.36e+03
848	4	13.8	558	2	S46326	intermediate filament	4.36e+03	921	4	13.8	858	1	ZPECPA	penicillin-binding pr	4.36e+03
849	4	13.8	559	2	S52950	period protein (clone	4.36e+03	922	4	13.8	858	2	A45613	protein 4.1, p4.1 - m	4.36e+03
850	4	13.8	561	2	E64980	hypothetical 62.1 kD	4.36e+03	923	4	13.8	862	2	S51493	major nitrogen regula	4.36e+03
851	4	13.8	562	2	S54464	probable membrane pro	4.36e+03	924	4	13.8	877	2	H64708	iron-regulated outer	4.36e+03
852	4	13.8	569	2	S74277	gene gp88 protein - m	4.36e+03	925	4	13.8	879	2	E71811	probable iron-regulat	4.36e+03
853	4	13.8	573	2	S06600	L-lactate dehydrogena	4.36e+03	926	4	13.8	880	2	B37443	protein-tyrosine kina	4.36e+03
854	4	13.8	577	1	S39804	moesin - plq	4.36e+03	927	4	13.8	883	1	RNBP17	DNA-directed RNA poly	4.36e+03
855	4	13.8	583	2	S38642	chaperonin 60 alpha c	4.36e+03	928	4	13.8	884	1	IJM5CE	E-cadherin precursor,	4.36e+03
856	4	13.8	585	2	F64159	hypothetical protein	4.36e+03	929	4	13.8	889	2	S26403	aconitate hydratase (	4.36e+03
857	4	13.8	586	2	S71235	chaperonin 60 alpha c	4.36e+03	930	4	13.8	899	2	S24223	uridyl transferase	4.36e+03
858	4	13.8	587	2	JX0278	5-aminolevulinate syn	4.36e+03	931	4	13.8	899	2	S38153	gene retII protein -	4.36e+03
859	4	13.8	587	2	S29783	pyruvate kinase (EC 2	4.36e+03	932	4	13.8	904	2	S53896	DNA mismatch repair p	4.36e+03
860	4	13.8	591	1	SYBYKT	lysine--tRNA ligase (	4.36e+03	933	4	13.8	907	1	VGBE7E	glycoprotein B precu	4.36e+03
861	4	13.8	591	2	S33542	catechol oxidase (EC	4.36e+03	934	4	13.8	915	2	S24577	ovarian protein B fru	4.36e+03
862	4	13.8	606	2	T03177	probable global trans	4.36e+03	935	4	13.8	928	1	SGBEMC	glycoprotein B precu	4.36e+03
863	4	13.8	609	1	A43458	replication protein A	4.36e+03	936	4	13.8	928	2	S64350	hypothetical protein	4.36e+03
864	4	13.8	610	2	G70178	exodeoxyribonuclease	4.36e+03	937	4	13.8	933	1	DEECOG	oxoglutarate dehydrog	4.36e+03
865	4	13.8	615	2	D44714	cell division protein	4.36e+03	938	4	13.8	933	2	S35233	methylene-tetrahydrof	4.36e+03
866	4	13.8	616	1	HNWZQD	hemagglutinin-neurami	4.36e+03	939	4	13.8	941	2	S78623	isoleucine--tRNA liga	4.36e+03
867	4	13.8	619	2	A84729	glucokinase regulator	4.36e+03	940	4	13.8	959	2	T03053	ribonucleoside-diphos	4.36e+03
868	4	13.8	619	2	C64568	primosomal protein re	4.36e+03	941	4	13.8	963	2	A53984	membrane alanyl amino	4.36e+03
869	4	13.8	626	2	C70104	DNA topoisomerase IV	4.36e+03	942	4	13.8	967	1	SYWTAT	alanine--tRNA ligase	4.36e+03
870	4	13.8	635	1	SYCHAL	5-aminolevulinate syn	4.36e+03	943	4	13.8	975	2	S33121	homeotic protein CDP	4.36e+03
871	4	13.8	637	2	A53125	restriction enzyme Bc	4.36e+03	944	4	13.8	988	2	I50611	protein-tyrosine kina	4.36e+03
872	4	13.8	642	2	S53434	plasma protein S prec	4.36e+03	945	4	13.8	994	1	MNVVM	nonstructural protein i	4.36e+03
873	4	13.8	644	1	FGHUA	fibrinogen alpha chai	4.36e+03	946	4	13.8	1000	1	A33620	myosin heavy chain I,	4.36e+03
874	4	13.8	662	1	A49882	histidine decarboxyla	4.36e+03	947	4	13.8	1023	2	A24639	Na+/K+-exchanging ATP	4.36e+03
875	4	13.8	663	2	S11521	CAMP-gated channel pr	4.36e+03	948	4	13.8	1025	1	DENCED	glutamate dehydrogena	4.36e+03
876	4	13.8	669	2	B42291	tail fiber protein H	4.36e+03	949	4	13.8	1055	2	JC5216	type I site-specific	4.36e+03
877	4	13.8	670	2	C71630	cytochrome C-type bio	4.36e+03	950	4	13.8	1058	2	JC1254	ubiquitin--protein li	4.36e+03
878	4	13.8	675	2	C64532	methyl-accepting chem	4.36e+03	951	4	13.8	1073	1	SYECCP	carbamoyl--phosphate s	4.36e+03
879	4	13.8	675	2	B71975	methyl-accepting chem	4.36e+03	952	4	13.8	1123	1	WMBEH7	UL37 protein - human	4.36e+03
880	4	13.8	676	1	WZBE43	gene 43 protein - hum	4.36e+03	953	4	13.8	1123	2	A39962	kinase-related transf	4.36e+03
881	4	13.8	682	1	A39666	protein kinase C (EC	4.36e+03	954	4	13.8	1136	1	S37845	protein-tyrosine kina	4.36e+03
882	4	13.8	686	2	G64618	cadmium-transporting	4.36e+03	955	4	13.8	1138	1	S24066	protein-tyrosine kina	4.36e+03
883	4	13.8	687	1	B32382	fcbb bifunctional pro	4.36e+03	956	4	13.8	1145	2	S55600	single-stranded DNA b	4.36e+03
884	4	13.8	689	2	S59797	hypothetical protein	4.36e+03	957	4	13.8	1169	2	JN0178	exodeoxyribonuclease	4.36e+03
885	4	13.8	691	2	S39867	competence protein co	4.36e+03	958	4	13.8	1176	2	JN0583	myosin-light-chain ki	4.36e+03
886	4	13.8	706	2	JQ2210	probable nucleic acid	4.36e+03	959	4	13.8	1208	2	S17286	period clock protein	4.36e+03
887	4	13.8	706	2	S19958	basic helix-loop-heli	4.36e+03	960	4	13.8	1243	2	S23393	Cad2--transporting ATP	4.36e+03
888	4	13.8	725	2	A59148	outer membrane protei	4.36e+03	961	4	13.8	1275	2	B28096	line-1 protein ORF2 -	4.36e+03
889	4	13.8	726	2	S73915	virulence-associated	4.36e+03	962	4	13.8	1278	2	B22671	TYB protein - yeast (	4.36e+03
890	4	13.8	737	1	S28942	protein kinase C (EC	4.36e+03	963	4	13.8	1280	2	B34087	hypothetical protein	4.36e+03
891	4	13.8	737	2	E36841	E2L protein - variola	4.36e+03	964	4	13.8	1296	1	HMSOLF	aggregation protein a	4.36e+03
892	4	13.8	739	2	E69794	ATP-dependent DNA hel	4.36e+03	965	4	13.8	1306	2	S24659	acetyl-CoA carboxylas	4.36e+03
893	4	13.8	740	2	E69420	hydrogenase expressio	4.36e+03	966	4	13.8	1323	1	RRWGM	RNA-directed RNA poly	4.36e+03
894	4	13.8	744	2	S68115	catalase (EC 1.11.1.6	4.36e+03	967	4	13.8	1324	1	VG1H59	E2 glycoprotein precu	4.36e+03
895	4	13.8	746	2	B32693	steroid receptor prot	4.36e+03	968	4	13.8	1328	2	B22999	TYB protein - yeast (	4.36e+03
896	4	13.8	747	2	S71478	homeotic protein Athb	4.36e+03	969	4	13.8	1378	2	I48751	protein-tyrosine kina	4.36e+03
897	4	13.8	751	1	SYBYMT	methionine--tRNA liga	4.36e+03	970	4	13.8	1404	2	A48196	protein-tyrosine kina	4.36e+03
898	4	13.8	754	2	G70124	hypothetical protein	4.36e+03	971	4	13.8	1405	1	DU2PA	DNA-directed DNA poly	4.36e+03
899	4	13.8	781	2	A49794	DNA topoisomerase (AT	4.36e+03	972	4	13.8	1506	2	JC5985	phosphoinositide 3-ki	4.36e+03



973 4 13.8 1572 2 S45251 SNF2alpha protein - h 4.36e+03  
974 4 13.8 1613 2 S39059 protein BRG1 - human 4.36e+03  
975 4 13.8 1670 1 CGHU3B collagen alpha 3(IV) 4.36e+03  
976 4 13.8 1736 2 A29176 sex-limited protein p 4.36e+03  
977 4 13.8 1755 2 S59975 Tyb protein - yeast ( 4.36e+03  
978 4 13.8 1755 2 S59866 Tyb protein - yeast ( 4.36e+03  
979 4 13.8 1755 2 S59845 Tyb protein - yeast ( 4.36e+03  
980 4 13.8 1755 2 S61763 Tyb protein - yeast ( 4.36e+03  
981 4 13.8 1755 2 S59949 Tyb protein - yeast ( 4.36e+03  
982 4 13.8 1755 2 S59955 Tyb protein - yeast ( 4.36e+03  
983 4 13.8 1756 2 S59983 Tyb protein - yeast ( 4.36e+03  
984 4 13.8 1770 2 S45842 Tyb protein - yeast ( 4.36e+03  
985 4 13.8 1776 1 RWPYM genome polyprotein - 4.36e+03  
986 4 13.8 1786 2 A57282 ankyrin-related prote 4.36e+03  
987 4 13.8 1863 2 S46217 protein-tyrosine-phos 4.36e+03  
988 4 13.8 1907 2 S50893 protein-tyrosine-phos 4.36e+03  
989 4 13.8 1955 1 AGCH agrin precursor - chi 4.36e+03  
990 4 13.8 1962 2 A32634 serine proteinase (EC 4.36e+03  
991 4 13.8 2048 1 ZLNZSE genome polyprotein - 4.36e+03  
992 4 13.8 2207 1 GNNY1P genome polyprotein (v 4.36e+03  
993 4 13.8 2319 2 A47004 coagulation factor VI 4.36e+03  
994 4 13.8 2492 1 C44213 nonstructural polypro 4.36e+03  
995 4 13.8 2629 2 I46569 apolipoprotein B - pi 4.36e+03  
996 4 13.8 2783 1 A41948 alpha-fetoprotein enh 4.36e+03  
997 4 13.8 3106 1 S33868 laminin alpha-2 chain 4.36e+03  
998 4 13.8 3164 1 WMBEH6 UL36 protein - human 4.36e+03  
999 4 13.8 3198 2 A43426 collagen alpha 2 fibr 4.36e+03  
1000 4 13.8 3259 1 A56539 giantin - human 4.36e+03

## ALIGNMENTS

RESULT 1  
ENTRY S29318 #type fragment  
TITLE lysophospholipase (EC 3.1.1.5) precursor - Penicillium notatum (fragment)  
ORGANISM #formal\_name Penicillium notatum  
DATE 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 13-Nov-1998  
ACCESSIONS S29318; S39881; S22045  
REFERENCE S29318  
#authors Masuda, N.; Kitamura, N.; Saito, K.  
#journal Eur. J. Biochem. (1991) 202:783-787  
#title Primary structure of protein moiety of Penicillium notatum phospholipase B deduced from the cDNA.  
#cross-references MUID:92111525  
#accession S29318  
#molecule\_type mRNA  
#residues 1-612 ##label MA1  
#cross-references EMBL:X60348; NID:g3183; PID:g3184  
#accession S39881  
#molecule\_type protein  
#residues 10-20185-199 ##label MA2  
CLASSIFICATION #superfamily yeast lysophospholipase  
KEYWORDS carboxylester hydrolase; glycoprotein  
FEATURE 1-9  
#domain signal sequence (fragment) #status predicted  
#label SIG  
#product lysophospholipase #status predicted #label MATV  
10-612  
41,81,116,223,267,  
306,335,427,440,  
446,477,498,526,  
532,567,571  
#binding\_site carbohydrate (Asn) (covalent) #status predicted  
#length 612 #checksum 7028  
SUMMARY  
Query Match 24.1%; Score 7; DB 2; Length 612;  
Best Local Similarity 100.0%; Pred. No. 1.29e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 481 PSPLVY 487  
15 PSPLVY 21

RESULT 2  
ENTRY S73603 #type complete  
TITLE hypothetical protein H03\_orf152 - Mycoplasma pneumoniae (ATCC 29342) (SGC3)  
ORGANISM #formal\_name Mycoplasma pneumoniae  
#variety ATCC 29342  
DATE 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 17-Jul-1998  
ACCESSIONS S73603  
REFERENCE S73327  
#authors Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
#journal Nucleic Acids Res. (1996) 24:4420-4449  
#title Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.  
#cross-references MUID:97105885  
#accession S73603  
#status preliminary: nucleic acid sequence not shown;  
#molecule\_type DNA  
#residues 1-152 ##label HIM  
#cross-references EMBL:AE000027; GB:U00089; NID:g1673941; PID:g1673948  
#note the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
GENETICS  
#genetic\_code SGC3  
SUMMARY #length 152 #molecular-weight 17463 #checksum 3357  
Query Match 20.7%; Score 6; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 6.61e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 20 PLVYVA 25  
17 PLVYVA 22  
RESULT 3  
ENTRY D69134 #type complete  
TITLE conserved hypothetical protein MTH272 - Methanobacterium thermoautotrophicum (strain Delta H)  
ORGANISM #formal\_name Methanobacterium thermoautotrophicum  
DATE 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 05-Jun-1998  
ACCESSIONS D69134  
REFERENCE A69000  
#authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrowski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
#journal J. Bacteriol. (1997) 179:7135-7155  
#title Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis and comparative genomics.  
#cross-references MUID:98037514  
#accession D69134  
#status preliminary: nucleic acid sequence not shown;  
#molecule\_type DNA  
#residues 1-205 ##label MTH  
#cross-references GB:AE000813; GB:AE000666; NID:g2621320; PID:g2621323  
#experimental\_source strain Delta H  
GENETICS  
#gene MTH272  
SUMMARY #length 205 #molecular-weight 22486 #checksum 8174



```

Query Match      20.7%; Score 6; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      146 IROGAS 151
Qy      24 IROGAS 29

RESULT 4
ENTRY  beta-lactamase (EC 3.5.2.6) precursor - Klebsiella pneumoniae
TITLE  #formal_name Klebsiella pneumoniae
ORGANISM 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change
DATE 18-Jun-1993
ACCESSIONS A24469
REFERENCE A24469
#authors Arakawa, Y.; Ohta, M.; Kido, N.; Fujii, Y.; Komatsu, T.;
#journal FEBS Lett. (1986) 207:69-74
#title Close evolutionary relationship between the chromosomally
#cross-references MUID:87030906 encoded beta-lactamase gene of Klebsiella pneumoniae and
#accession A24469 the TEM beta-lactamase gene mediated by R plasmids.
#molecule_type DNA
#residues 1-279 ##label ARA
CLASSIFICATION #superfamily beta-lactamase I
KEYWORDS antibiotic resistance; hydrolase
SUMMARY #length 279 #molecular-weight 30287 #checksum 5625

Query Match      20.7%; Score 6; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      16 PLVVYA 21
Qy      17 PLVVYA 22

RESULT 5
ENTRY  cytochrome aa3 quinol oxidase (subunit II) qoxA - Bacillus
TITLE  subtilis
ALTERNATE_NAMES qoxA
ORGANISM Bacillus subtilis
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
17-Mar-1999
ACCESSIONS E69687; A38129; S39692
REFERENCE A69380
#authors Kunat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.A.; Codani, J.F.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.K.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fahret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kasahara, Y.; Klaer-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maeel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott,
A.M.; Presecan, E.; Pujic, P.; Furnelle, B.; Rapoport, G.;

```

```

Key, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Toynoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenecker, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references MUID:98044033
#accession E69687
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-321 ##label KUN
#cross-references GB:299123; GB:AL009126; NID:g2636240; PID:e1186316;
PID:g2636352
#experimental_source strain 168
REFERENCE A38129
#authors Santana, M.; Kunst, F.; Hullo, M.F.; Rapoport, G.; Danchin,
A.; Glaser, P.
J. Biol. Chem. (1992) 267:10225-10231
Molecular cloning, sequencing, and physiological
characterization of the qox operon from Bacillus subtilis
encoding the aa3-600 quinol oxidase.
#cross-references MUID:92268053
#accession A38129
#molecule_type DNA
#residues 'MHFRFTISIGRMV', 2-38, 'D', 39-321 ##label SAN
#cross-references GB:M85548; NID:gl43395; PID:gl43396
#note sequence extracted from NCBI backbone (NCBIN:103632,
NCBIP:103599)
REFERENCE S39655
#authors Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales,
W.; Hullo, M.F.; Ionescu, M.; Lubochinsky, B.; Marcelino,
L.; Moszer, I.; Presecan, E.; Santana, M.; Schneider, E.;
Schweizer, J.; Vertes, A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. (1993) 10:371-384
Bacillus subtilis genome project: cloning and sequencing of
the 97 kb region from 325 degrees to 333 degrees.
#cross-references MUID:95020537
#accession S39692
#molecule_type DNA
#residues 'V', 2-38, 'D', 39-321 ##label GLA
#cross-references EMBL:X73124
GENETICS qoxA
#gene
CLASSIFICATION #superfamily bo-type ubiquinol oxidase chain II precursor;
cytochrome-c oxidase chain II homology
FEATURE 31-229 #domain cytochrome-c oxidase chain II homology #label
CO2
SUMMARY #length 321 #molecular-weight 36253 #checksum 2559
Query Match      20.7%; Score 6; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      130 PLVVYA 135
Qy      17 PLVVYA 22

RESULT 6
ENTRY  #type complete
TITLE  ribosomal protein L4.e.B, cytosolic - yeast (Saccharomyces
cerevisiae)
ALTERNATE_NAMES S50993
protein D3221; protein p2B362; protein yD8119.17; protein

```

YDR012w; ribosomal protein YL2.B  
#formal\_name Saccharomyces cerevisiae  
#sequence\_revision 11-Aug-1995 #text\_change  
11-Feb-1995 #sequence\_revision 11-Aug-1995 #text\_change  
24-Sep-1998  
#accessions S50993; S63419; S42260; S67825; S72110  
#reference S50976  
#authors Murphy, L.; Richards, C.; Gentles, S.; Harris, D.  
#submission submitted to the EMBL Data Library, January 1995  
#accession S50993  
#molecule\_type DNA  
#residues 1-362 #label MUR  
#cross-references EMBL:248008; NID:g642799; PID:g642817  
#reference S63416  
#authors Eide, L.G.; Sander, C.; Prydz, H.  
#submission submitted to the EMBL Data Library, February 1996  
#description Sequencing and analysis of a 35.4 kb region on the left arm  
of chromosome IV for Saccharomyces cerevisiae reveal 23  
open reading frames.  
#accession S63419  
#molecule\_type DNA  
#residues 1-362 #label EID  
#cross-references EMBL:X95966; NID:g1216215; PID:e225618; PID:g1216219  
#reference S42260  
#authors Presutti, C.  
#submission submitted to the EMBL Data Library, February 1989  
#accession S42260  
#molecule\_type DNA  
#residues 1-87, 'R', '89-200 #label PRE  
#cross-references EMBL:M22583; NID:g172414; PID:g553139  
#reference S67822  
#authors Prydz, H.; Eide, L.G.  
#submission submitted to the Protein Sequence Database, July 1996  
#accession S67825  
#molecule\_type DNA  
#residues 1-362 #label PRY  
#cross-references EMBL:274308; NID:g1431431; PID:e253289; PID:g1431432;  
#cross-references EMBL:274308; NID:g1431431; PID:e253289; PID:g1431432;  
#experimental\_source strain S288C  
#reference S72107  
#authors Eide, L.G.; Sander, C.; Prydz, H.  
#journal Yeast (1996) 12:1085-1090  
#title Sequencing and analysis of a 35.4 kb region on the left arm  
of chromosome IV from Saccharomyces cerevisiae reveal 23  
open reading frames.  
#cross-references MUID:97051598  
#accession S72110  
#status nucleic acid sequence not shown; translation not shown  
#molecule\_type DNA  
#residues 1-362 #label EIW  
#cross-references EMBL:X95966; NID:g1216215; PID:e225618; PID:g1216219  
#note the nucleotide sequence was submitted to the EMBL Data  
Library, February 1996  
GENETICS  
#gene SGD:RPL2B  
#cross-references MIPS:YDR012w; SGD:S0002419  
#map\_position 4R  
CLASSIFICATION  
#superfamily rat ribosomal protein L4  
KEYWORDS  
#cytosol; protein biosynthesis; ribosome  
#length 362 #molecular\_weight 39062 #checksum 6572  
SUMMARY  
Query Match 20.7%; Score 6; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 6.61e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 205 PLVVYA 210  
| | | | |  
QY 17 PLVVYA 22  
RESULT 7  
ENTRY #type complete  
TITLE ribosomal protein L4.e.A, cytosolic - yeast (Saccharomyces  
cerevisiae)

protein YBR0315; protein YBR031w; ribosomal protein rp2;  
ribosomal protein YL2.A  
#formal\_name Saccharomyces cerevisiae  
#sequence\_revision 09-Sep-1994 #text\_change  
26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change  
06-Feb-1998  
#accessions S45887; S46561; A28656; S45500  
#reference S45875  
#authors Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.  
#submission submitted to the Protein Sequence Database, August 1994  
#accession S45887  
#molecule\_type DNA  
#residues 1-362 #label GRI  
#cross-references EMBL:Z35900; NID:g536243; PID:g536244; MIPS:YBR031w  
#experimental\_source strain S288C  
#reference S46551  
#authors Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.  
#journal Yeast (1994) 10(Suppl.A):S75-S80  
#title The complete sequence of a 33 kb fragment on the right arm of  
chromosome II from Saccharomyces cerevisiae reveals 16 open  
reading frames, including ten new open reading frames, five  
previously identified genes and a homologue of the SCO1  
gene.  
#accession S46561  
#status nucleic acid sequence not shown; translation not shown  
#molecule\_type DNA  
#residues 1-362 #label SMI  
#cross-references EMBL:X76078; NID:g498748; PID:g498759  
#experimental\_source strain S288C  
#note the nucleotide sequence was submitted to the EMBL Data  
Library, November 1993  
#reference A28656  
#authors Presutti, C.; Lucifoli, A.; Bozzoni, I.  
#journal J. Biol. Chem. (1988) 263:6188-6192  
#title Ribosomal protein L2 in Saccharomyces cerevisiae is  
homologous to ribosomal protein L1 in Xenopus laevis.  
#cross-references MUID:88198160  
#accession A28656  
#molecule\_type DNA  
#residues 1-37, 'L', '39-143, 'T', '145-156, 'D', '158-223, 'S', '225-240, 'S',  
#cross-references GB:J03195; NID:g172412; PID:g172413  
#note the authors translated the codon CAA for residue 59 as  
Glu and CAC for residue 311 as Lys  
#reference S45500  
#authors Takahara, H.; Tsunasawa, S.; Miyagi, M.; Warner, J.R.  
#journal J. Biol. Chem. (1992) 267:5442-5445  
#title NH2-terminal acetylation of ribosomal proteins of  
Saccharomyces cerevisiae.  
#accession S45500  
#molecule\_type protein  
#residues 2-21 #label TAK  
GENETICS  
#gene SGD:RPL2A  
#cross-references SGD:S0000235; MIPS:YBR031w  
#map\_position 2R  
CLASSIFICATION  
#superfamily rat ribosomal protein L4  
KEYWORDS  
#acetylated amino end; blocked amino end; cytosol; protein  
biosynthesis; ribosome  
FEATURE  
2-362  
#product ribosomal protein L4.e #status experimental  
#label MAT\N  
#modified\_site acetylated amino end (Ser) (in mature  
form) #status experimental  
SUMMARY  
#length 362 #molecular\_weight 39092 #checksum 6838  
Query Match 20.7%; Score 6; DB 2; Length 362;  
Best Local Similarity 100.0%; Pred. No. 6.61e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 205 PLVVYA 210  
| | | | |  
QY 17 PLVVYA 22

```

      8
RESULT      S61566      #type complete
ENTRY       BP1 protein - yeast (Saccharomyces cerevisiae)
TITLE       protein P0745; protein YPL255w
ALTERNATE_NAMES
ORGANISM     #formal_name Saccharomyces cerevisiae
DATE         09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change
              06-Feb-1998
ACCESSIONS  S61566; S65286; S65284
REFERENCE    S61566
#authors     Xue, Z.; Shan, X.; Melese, T.
#submission  submitted to the EMBL Data Library, October 1995
#accession   S61566
#molecule_type DNA
#residues    1-385 #label XUE
#cross-references EMBL:X92658; NID:g1113128; PID:e209716; PID:g1113129
#experimental_source strain S288C
REFERENCE    S64935
#authors     Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.
#submission  submitted to the Protein Sequence Database, May 1996
#accession   S65286
#molecule_type DNA
#residues    1-385 #label MES
#cross-references EMBL:Z73611; NID:g1370523; PID:e247082; PID:g1370524;
              MIPS:YPL255w
#experimental_source strain S288C (AB972)
REFERENCE    S64899
#authors     Pohl, T.M.
#submission  submitted to the Protein Sequence Database, May 1996
#accession   S65284
#molecule_type DNA
#residues    298-385 #label POH
#cross-references EMBL:Z73611; MIPS:YPL255w
#experimental_source strain S288C (AB972)
GENETICS
#gene        SGD:BBP1
#map_position 16L
#cross-references SGD:S0006176; MIPS:YPL255w
SUMMARY      #length 385 #molecular-weight 45384 #checksum 9921

Query Match      20.7%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 191 KPSPPL 196
QY 13 KPSPPL 18

RESULT      9
ENTRY       A25351      #type complete
TITLE       ubiquinol--cytochrome-c reductase (EC 1.10.2.2) 44k core
              protein precursor - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES
              coenzyme QH2-cytochrome c reductase 44k core protein; protein
              YBL0403; protein YBL045c; type 1 dehydrogenase
ORGANISM     #formal_name Saccharomyces cerevisiae
DATE         30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change
              17-Mar-1999
ACCESSIONS  A25351; S50286; S45779; S42500
REFERENCE    A25351
#authors     Tzagoloff, A.; Wu, M.; Crivellone, M.
#journal     J. Biol. Chem. (1986) 261:17163-17169
#title       Assembly of the mitochondrial membrane system.
              Characterization of COR1, the structural gene for the
              44-kilodalton core protein of yeast coenzyme QH2-cytochrome
              c reductase.
#cross-references MUID:87057439
#accession   A25351
#molecule_type DNA
#residues    1-457 #label TZ
#cross-references EMBL:J02636; NID:g171255; PID:g171256
REFERENCE    S50284

#authors     de Mergifosse, P.; Jacques, B.; Jonniaux, J.L.; Purnelle, B.;
              Skala, J.; Goffeau, A.
#journal     Yeast (1994) 10:1489-1496
#title       The sequence of a 22.4 kb DNA fragment from the left arm of
              yeast chromosome II reveals homologues to bacterial proline
              synthetase and murine alpha-adaptin, as well as a new
              permease and a DNA-binding protein.
#cross-references MUID:95176707
#accession   S50286
#status      nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues    1-457 #label DEF
#cross-references EMBL:X78214; NID:g463261; PID:g463264
#note        the nucleotide sequence was submitted to the EMBL Data
              Library, March 1994
REFERENCE    S45745
#authors     Goffeau, A.; Jonniaux, J.L.; Purnelle, B.; Skala, J.; de
              Mergifosse, P.; van Dyck, L.
#submission  submitted to the Protein Sequence Database, August 1994
#accession   S45779
#molecule_type DNA
#residues    1-457 #label GOF
#cross-references EMBL:Z35806; NID:g536064; PID:g536065; MIPS:YBL045c
GENETICS
#gene        SGD:COR1
#map_position 2L
#cross-references SGD:S0000141; MIPS:YBL045c
#genome      nuclear
CLASSIFICATION #superfamily mitochondrial processing peptidase alpha chain
KEYWORDS       mitochondrion; oxidoreductase; respiratory chain
FEATURE        1-26
              #domain transit peptide (mitochondrion) #status
              predicted #label TNP\
              #product ubiquinol--cytochrome-c reductase 44k core
              protein #status predicted #label MAT
              #length 457 #molecular-weight 50227 #checksum 101
SUMMARY

Query Match      20.7%; Score 6; DB 2; Length 457;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 TVTSKT 9
QY 2 TVTSKT 7

RESULT      10
ENTRY       POLJEV      #type complete
TITLE       gag polyprotein - equine infectious anemia virus
ALTERNATE_NAMES
              core polyprotein
              core protein 9; core protein p11; core protein p15; core
              protein p26
CONTAINS      #formal_name equine infectious anemia virus
              host Equus caballus (domestic horse)
              28-May-1986 #sequence_revision 28-May-1986 #text_change
              26-Feb-1999
ACCESSIONS  A03949; A27842; A41991
REFERENCE    A03949
#authors     Stephens, R.M.; Casey, J.W.; Rice, N.R.
#journal     Science (1986) 231:589-594
#title       Equine infectious anemia virus gag and pol genes: relatedness
              to visna and AIDS virus.
#cross-references MUID:86122873
#accession   A03949
#molecule_type DNA
#residues    1-486 #label STE
#cross-references GB:M16575; GB:K03334; GB:M11337; GB:M14855;
              NID:g323836; PID:g323837
REFERENCE    A27842
#authors     Kawakami, T.; Sherman, L.; Dahlberg, J.; Gazit, A.; Yaniv,
              A.; Tronick, S.R.; Aaronson, S.A.
#journal     Virology (1987) 158:300-312
#title       Nucleotide sequence analysis of equine infectious anemia
```

```

virus proviral DNA.
#cross-references MUID:87236196
#accession A27842
#molecule_type DNA
#residues 1-486 #label KAW
#cross-references GB:M16575; GB:K03334; GB:M11337; GB:M14855;
NID:g323836; PID:g323837
##experimental_source clone 1369
REFERENCE
A41991
#authors Perry, S.T.; Flaherty, M.T.; Kelley, M.J.; Clabough, D.L.;
Tronick, S.R.; Coggins, L.; Whetter, L.; Lengel, C.R.;
Fuller, F.
#journal J. Virol. (1992) 66:4085-4097
#title The surface envelope protein gene region of equine infectious
anemia virus is not an important determinant of tropism in
vitro.
#cross-references MUID:92292230
#accession A41991
#molecule_type DNA
#residues 1-486 #label PER
#cross-references GB:M87581; NID:g290627; PID:g290628
#experimental_source strain CL22
GENETICS
#gene gag
#classification #superfamily AIDS-related virus gag polyprotein
#keywords core protein; polyprotein
FEATURE
1-124 #product core protein p15 #status predicted #label p15\
125-359 #product core protein p26 #status predicted #label p26\
360-435 #product core protein p11 #status predicted #label p11\
436-486 #product core protein p9 #status predicted #label p09
SUMMARY
#length 486 #molecular-weight 54808 #checksum 3630
Query Match 20.7%; Score 6; DB 1; Length 486;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 276 NIROGA 281
|||||
Qy 23 NIROGA 28
RESULT 11
ENTRY #type complete
TITLE integral membrane protein HP0228 homolog 1 - Synechocystis
ALTERNATE_NAMES SP. (strain PCC 6803)
ORGANISM protein sir1229
#variety #formal_name Synechocystis sp.
DATE 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
21-Aug-1998
ACCESSIONS S77411
REFERENCE S74322
#authors Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;
Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo,
S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
Yasuda, M.; Tabata, S.
#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular
cyanobacterium Synechocystis sp. PCC6803. II. Sequence
determination of the entire genome and assignment of
potential protein-coding regions.
#cross-references MUID:97061201
#accession S77411
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-498 #label KAN
#cross-references EMBL:P09096; GB:AB001339; NID:g1652492; PID:d1018247;
PID:g1652593
#note the nucleotide sequence was submitted to the EMBL Data
Library, June 1996
CLASSIFICATION #superfamily integral membrane protein HP0228
SUMMARY #length 498 #molecular-weight 54495 #checksum 5215
Query Match 20.7%; Score 6; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 173 FPSPLV 178
|||||
Qy 14 FPSPLV 19
RESULT 12
ENTRY #type complete
TITLE hypothetical protein F411.7 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02382
REFERENCE Z14177
#authors Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.;
Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.;
Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter,
J.C.
#submission submitted to the EMBL Data Library, May 1998
#description Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
#accession T02382
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-583 #label ROU
#cross-references EMBL:AC004521; NID:g3128166; PID:g3128172
GENETICS
#map_position 2
#note F411.7
SUMMARY #length 583 #molecular-weight 64255 #checksum 1508
Query Match 20.7%; Score 6; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 30 KPSPPL 35
|||||
Qy 13 KPSPPL 18
RESULT 13
ENTRY #type complete
TITLE rab geranylgeranyl transferase component A - yeast
(Saccharomyces cerevisiae)
ALTERNATE_NAMES protein O6676; protein YOR370c
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
06-Feb-1998
ACCESSIONS S47917; S48222; S27431; S67282; S43545
REFERENCE S47917
#authors Fujimura, K.; Tanaka, K.; Nakano, A.; Tche, A.
#journal J. Biol. Chem. (1994) 269:9205-9212
#title The Saccharomyces cerevisiae MS14 gene encodes the yeast
counterpart of component A of Rab
geranylgeranyltransferase.
#cross-references MUID:94179341
#accession S47917
#molecule_type DNA
#residues 1-603 #label FUJ
#cross-references EMBL:D26441; NID:g452240; PID:d1006002; PID:g452241
S48222
REFERENCE
#authors Ragnini, A.; Teply, R.; Waldherr, M.; Voskova, A.; Schweyen,
R.J.
#journal Curr. Genet. (1994) 26:308-314
#title The yeast protein Mps6p, a homologue of the rabGDI and human
cholesterolemia proteins, affects cytoplasmic and
mitochondrial functions.

```

```
#cross-references MUID:95188267
#accession S48222
  #molecule_type DNA
  #residues 1-299,'P',301-603 ##label RAG
  #cross-references EMBL:X70339; NID:g473150; PID:g473151
REFERENCE S27431
#authors Waldherr, M.; Voskova, A.; Schweyen, R.J.
#submission submitted to the EMBL Data Library, April 1992
#accession S27431
  #molecule_type DNA
  #residues 'MIPPRV',138-299,'P',301-594 ##label WAL
  #cross-references EMBL:M90844; NID:g171989; PID:g171990
REFERENCE S67261
#authors Dellius, H.; Hebling, U.; Hofmann, B.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67262
  #molecule_type DNA
  #residues 1-603 ##label DEL
  #cross-references EMBL:Z75278; NID:g1420799; PID:e252194; PID:g1420800;
  MIPS:YOR370c
##experimental_source strain S288c
GENETICS
#gene SGD:MRS6; MSI4
##cross-references SGD:S0005897; MIPS:YOR370c
#map_position 15R
KEYWORDS transmembrane protein
FEATURE
49-65 #domain transmembrane #status predicted #label TMM
SUMMARY #length 603 #molecular-weight 67373 #checksum 3038

Query Match 20.7%; Score 6; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 PSLVV 29
QY 15 PSLVV 20

RESULT 14
ENTRY B34469 #type complete
TITLE pullulanase secretion protein pULD precursor - Klebsiella
  pneumoniae (strain UNF5023)
ORGANISM #formal_name Klebsiella pneumoniae
DATE 15-Jun-1990 #sequence_revision 31-Dec-1993 #text_change
20-Feb-1995
ACCESSIONS B34469; B31394
REFERENCE A34469
#authors d'Enfert, C.; Reys, I.; Wandersman, C.; Pugsley, A.P.
#journal J. Biol. Chem. (1989) 264:17462-17468
#title Protein secretion by gram-negative bacteria. Characterization
  of two membrane proteins required for pullulanase secretion
  by Escherichia coli K-12.
#cross-references MUID:90008916
#accession B34469
  #molecule_type DNA
  #residues 1-660 ##label DEN
  #cross-references GB:M32613
GENETICS
#gene pULD
#start_codon TTG
SUMMARY #length 660 #molecular-weight 70658 #checksum 1276

Query Match 20.7%; Score 6; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 215 NKDTSK 220
QY 8 NKDTSK 13

RESULT 15
```

```
ENTRY JG0168 #type complete
TITLE gob-5 protein - Mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 18-Mar-1999 #sequence_revision 18-Mar-1999 #text_change
18-Mar-1999
ACCESSIONS JG0168
REFERENCE JG0168
#authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351
#title cloning and identification of the gene gob-5, which is
  expressed in intestinal goblet cells in mice.
#accession JG0168
#status preliminary
#residues 1-913 ##label KOM
##cross-references DDBJ:AB016592
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 20.7%; Score 6; DB 3; Length 913;
Best Local Similarity 100.0%; Pred. No. 6.61e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 615 IRQGAS 620
QY 24 IRQGAS 29

Search completed: Sat Aug 28 14:58:53 1999
Job time : 47 secs.
```

**This Page Blank (uspto)**

\*\*\*\*\*  
MISRA  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:55:57 1999; MasPar time 3.44 Seconds  
Tabular output not generated. 238.520 Million cell updates/sec

Title: >US-09-049-696-46  
Description: (1-29) from US09049696.pep  
Perfect Score: 29  
Sequence: 1 ITVTSKTKDTSKFPSPLVVYNIQAS 29

Scoring table: TABLE uniprotatable  
Gap 60

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: swissprot37  
1:swissprot

Statistics: Mean 2.774; Variance 0.397; scale 6.982

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	7	24.1	612	1	PLBI PENNO LYSOPHOSPHOLIPASE PREC	4.11e-02
2	6	20.7	279	1	BLAC KLEPN BETA-LACTAMASE PRECURS	2.87e+00
3	6	20.7	323	1	QOX2_BACSU QUINOL OXIDASE POLYPEP	2.87e+00
4	6	20.7	361	1	RL4E_YEAST 60S RIBOSOMAL PROTEIN	2.87e+00
5	6	20.7	361	1	RL4E_YEAST 60S RIBOSOMAL PROTEIN	2.87e+00
6	6	20.7	385	1	BBP1_YEAST BBP1 PROTEIN	2.87e+00
7	6	20.7	457	1	UCR1_YEAST UBIQUINOL-CYTOCHROME-C	2.87e+00
8	6	20.7	457	1	PURA_CAEEL PROBABLE ADENYLOSUCCIN	2.87e+00
9	6	20.7	486	1	GAG_EIAYI GAG POLYPROTEIN [CONTA	2.87e+00
10	6	20.7	603	1	RAEP_YEAST RAB PROTEINS GERANYLGE	2.87e+00
11	6	20.7	660	1	GSPD_KLEPN GENERAL SECRETION PATH	2.87e+00
12	6	20.7	832	1	VGLB_HSV7J GLYCOPROTEIN B PRECURS	2.87e+00
13	6	20.7	1857	1	PAS2_PENPA FATTY ACID SYNTHASE, S	2.87e+00
14	5	17.2	41	1	PSBE_EUGGR CYTOCHROME B559 BETA C	1.25e+02
15	5	17.2	85	1	PTPH_HAEIN PHOSPHOCARRIER PROTEIN	1.25e+02
16	5	17.2	85	1	PTPH_HAEIN PHOSPHOCARRIER PROTEIN	1.25e+02
17	5	17.2	85	1	PTPH_HAEIN PHOSPHOCARRIER PROTEIN	1.25e+02
18	5	17.2	98	1	NULM_PHOVI NADH-UBIQUINONE OXIDOR	1.25e+02
19	5	17.2	105	1	SM52_ICTPU SOMATOSTATIN II PRECUR	1.25e+02
20	5	17.2	105	1	MTL1_TETPI METALLOTHIONEIN (MT-1)	1.25e+02
21	5	17.2	108	1	VG01_VZVD HYPOTHETICAL GENE 1 PR	1.25e+02
22	5	17.2	119	1	HV2C_HUMAN IG HEAVY CHAIN V-II RE	1.25e+02
23	5	17.2	125	1	FAB2_AMBME FATTY ACID-BINDING PRO	1.25e+02

FATTY ACID-BINDING PRO 1.25e+02  
FATTY ACID-BINDING PRO 1.25e+02  
60S RIBOSOMAL PROTEIN 1.25e+02  
60S RIBOSOMAL PROTEIN 1.25e+02  
60S RIBOSOMAL PROTEIN 1.25e+02  
BIOPOLYMER TRANSPORT E 1.25e+02  
EARLY PROTEIN GP17 1.25e+02  
AMICYANIN PRECURSOR 1.25e+02  
OPACITY ASSOCIATED PRO 1.25e+02  
LYSOZYME PRECURSOR (EC 1.25e+02  
50S RIBOSOMAL PROTEIN 1.25e+02  
IG HEAVY CHAIN PRECURS 1.25e+02  
HYPOTHETICAL PROTEIN M 1.25e+02  
HYPOTHETICAL PROTEIN I 1.25e+02  
NITROGEN REGULATORY II 1.25e+02  
HYPOTHETICAL 18.1 KD P 1.25e+02  
PHOTOSYSTEM I REACTION 1.25e+02  
NOCICEPTIN PRECURSOR ( 1.25e+02  
PHOSPHORIBOSYLAMINOIM 1.25e+02  
RIBULOSE BIPHOSPHATE 1.25e+02  
CELL DIVISION CONTROL 1.25e+02  
SPORULATION-SPECIFIC P 1.25e+02  
HYPOTHETICAL 20.8 KD P 1.25e+02  
COLD-REGULATED PROTEIN 1.25e+02  
FIBROLASE (EC 3.4.24.7 1.25e+02  
GENOME POLYPROTEIN [CO 1.25e+02  
SUPEROXIDE DISMUTASE I 1.25e+02  
HYPOTHETICAL 24.2 KD P 1.25e+02  
TAIL FIBER PROTEIN GP3 1.25e+02  
60S RIBOSOMAL PROTEIN 1.25e+02  
EPENDYMIN PRECURSOR (E 1.25e+02  
HYPOTHETICAL 24.6 KD P 1.25e+02  
25 KD BRAIN-SPECIFIC P 1.25e+02  
GENERAL SECRETION PATH 1.25e+02  
HOMEOBOX PROTEIN MSH-D 1.25e+02  
PROBABLE HELICASE (ORF 1.25e+02  
HYPOTHETICAL 26.3 KD H 1.25e+02  
PEPTIDYL-PROLYL CIS-TR 1.25e+02  
CHLOROPLAST SMALL HEAT 1.25e+02  
CHLOROPLAST SMALL HEAT 1.25e+02  
TRIOSEPHOSPHATE ISOMER 1.25e+02  
COMPLEMENT C1Q SUBCOMP 1.25e+02  
60S RIBOSOMAL PROTEIN 1.25e+02  
CYTOCHROME C OXIDASE P 1.25e+02  
TRANSCRIPTION FACTOR A 1.25e+02  
PROBABLE IRON CHELATIN 1.25e+02  
BRI PROTEIN (PUTATIVE 1.25e+02  
OPACITY-RELATED PROTEI 1.25e+02  
GLUTAMINE-BINDING PROT 1.25e+02  
LYSOZYME (EC 3.2.1.17) 1.25e+02  
NORO PROTEIN 1.25e+02  
PYRROLINE-5-CARBOXYLAT 1.25e+02  
PEP11 PROTEIN 1.25e+02  
SUCCINOGLYCAN BIOSYNTH 1.25e+02  
HYPOTHETICAL 32.5 KD P 1.25e+02  
CHORION PROTEIN S36 1.25e+02  
CHORION PROTEIN S36 1.25e+02  
GENOME POLYPROTEIN [CO 1.25e+02  
GENOME POLYPROTEIN [CO 1.25e+02  
RIBONUCLEOSIDE-DIPHOS 1.25e+02  
HOMEOBOX PROTEIN GSH-2 1.25e+02  
PROBABLE QUINOLINATE S 1.25e+02  
HYPOTHETICAL FRANScrip 1.25e+02  
PROTEIN N-TERMINAL ASP 1.25e+02  
CARBONIC ANHYDRASE IV 1.25e+02  
HYPOTHETICAL 33.4 KD P 1.25e+02  
HYPOTHETICAL 34.2 KD P 1.25e+02  
HYPOTHETICAL 35.9 KD P 1.25e+02  
D7\_AEDAE D7 PROTEIN PRECURSOR 1.25e+02  
60S ACIDIC RIBOSOMAL P 1.25e+02  
HYPOTHETICAL PROTEIN M 1.25e+02

97	5	17.2	328	1	IPNS_NOCOLA	ISOPENICILLIN N SYNTHET	1.25e+02	170	5	17.2	505	1	ALAT_SCHPO	PUTATIVE ALANINE AMINO	1.25e+02
98	5	17.2	331	1	SRAL_CAEEL	SRA-1 PROTEIN.	1.25e+02	171	5	17.2	510	1	HEG2_HALPO	PROTEIN HGV2.	1.25e+02
99	5	17.2	332	1	MC4R_HUMAN	MELANOCORTIN-4 RECEPTO	1.25e+02	172	5	17.2	511	1	AYRB_XENIA	ACTIVIN RECEPTOR TYPE	1.25e+02
100	5	17.2	332	1	MC4R_RAT	MELANOCORTIN-4 RECEPTO	1.25e+02	173	5	17.2	512	1	AYRB_HUMAN	ACTIVIN RECEPTOR TYPE	1.25e+02
101	5	17.2	333	1	Y720_METJA	HYPOTHETICAL PROTEIN M	1.25e+02	174	5	17.2	512	1	NARH_ECOLI	RESPIRATORY NITRATE RE	1.25e+02
102	5	17.2	339	1	LSP1_HUMAN	LYMPHOCYTE-SPECIFIC PR	1.25e+02	175	5	17.2	513	1	V1L1_PAPVO	MAJOR CAPSID PROTEIN L	1.25e+02
103	5	17.2	341	1	NU2M_DROYA	NADH-UBIQUINONE OXIDOR	1.25e+02	176	5	17.2	518	1	TYD1_PAPSO	TYROSINE/DOPA DECARBOX	1.25e+02
104	5	17.2	341	1	NU2M_DROME	NADH-UBIQUINONE OXIDOR	1.25e+02	177	5	17.2	520	1	TR14_FUSSP	TRICHODIENE OXYGENASE	1.25e+02
105	5	17.2	342	1	CAC4_CHICK	UMP-N-ACETYLNEURAMINAT	1.25e+02	178	5	17.2	523	1	TYD5_PAPSO	TYROSINE/DOPA DECARBOX	1.25e+02
106	5	17.2	347	1	NADA_ECOLI	QUINOLINATE SYNTHETASE	1.25e+02	179	5	17.2	527	1	LAC5_TRAVI	LACASE 5 PRECURSOR (E	1.25e+02
107	5	17.2	353	1	41KD_LACHE	41 KD PROTEIN.	1.25e+02	180	5	17.2	527	1	LAC5_TRAVE	LACASE PRECURSOR (EC	1.25e+02
108	5	17.2	355	1	VALL1_BMWV	ALL PROTEIN.	1.25e+02	181	5	17.2	536	1	AYRB_MOUSE	ACTIVIN RECEPTOR TYPE	1.25e+02
109	5	17.2	363	1	YK57_YEAST	HYPOTHETICAL 39.9 KD P	1.25e+02	182	5	17.2	538	1	LSHR_SHEEP	LUTROPIN-CHORIOGNADOT	1.25e+02
110	5	17.2	365	1	NADA_SALTY	QUINOLINATE SYNTHETASE	1.25e+02	183	5	17.2	546	1	SR8_MOUSE	SRC SUBSTRATE CORTACTI	1.25e+02
111	5	17.2	370	1	WNT1_BRARE	WNT-1 PROTEIN PRECURSOR	1.25e+02	184	5	17.2	547	1	ILVB_MYCTU	PROBABLE ACETOLACTATE	1.25e+02
112	5	17.2	371	1	HYAL_BPH44	HYALURONGLUCOSAMINIDA	1.25e+02	185	5	17.2	550	1	INV1_HUMAN	INVERTASE PRECURSOR (E	1.25e+02
113	5	17.2	373	1	LEU3_CANAL	3-ISOPROPYLMALATE DEHY	1.25e+02	186	5	17.2	550	1	SR8_HUMAN	SRC SUBSTRATE CORTACTI	1.25e+02
114	5	17.2	373	1	LEU3_CANMA	3-ISOPROPYLMALATE DEHY	1.25e+02	187	5	17.2	551	1	GZF3_YEAST	GZF3 PROTEIN.	1.25e+02
115	5	17.2	374	1	HS83_DROVI	HEAT SHOCK PROTEIN 83	1.25e+02	188	5	17.2	553	1	DMR9_HUMAN	DMR-N9 PROTEIN (PROTEI	1.25e+02
116	5	17.2	375	1	HS83_DROPS	HEAT SHOCK PROTEIN 83	1.25e+02	189	5	17.2	557	1	FM04_HUMAN	DIMETHYLANILINE MONOX	1.25e+02
117	5	17.2	375	1	HS83_DROSI	HEAT SHOCK PROTEIN 83	1.25e+02	190	5	17.2	557	1	Y423_MYCPN	HYPOTHETICAL PROTEIN M	1.25e+02
118	5	17.2	376	1	KITH_HSV23	THYMIDINE KINASE (EC 2	1.25e+02	191	5	17.2	561	1	VGLF_SENDS	FUSION GLYCOPROTEIN PR	1.25e+02
119	5	17.2	378	1	NAGA_VIBCH	N-ACETYLGLUCOSAMINE-6-	1.25e+02	192	5	17.2	565	1	VGLF_SENDS	FUSION GLYCOPROTEIN PR	1.25e+02
120	5	17.2	379	1	LYXA_ASPNG	PECTIN LYASE A PRECURS	1.25e+02	193	5	17.2	567	1	CH13_CANAL	CHITINASE 3 PRECURSOR	1.25e+02
121	5	17.2	380	1	LEU3_PHACH	3-ISOPROPYLMALATE DEHY	1.25e+02	194	5	17.2	568	1	PAP_YEAST	POLY(A) POLYMERASE (EC	1.25e+02
122	5	17.2	381	1	YBD2_YEAST	HYPOTHETICAL 41.7 KD P	1.25e+02	195	5	17.2	572	1	DPY2_BOVIN	DIHYDROPYRIMIDINASE RE	1.25e+02
123	5	17.2	382	1	NAGA_ECOLI	N-ACETYLGLUCOSAMINE-6-	1.25e+02	196	5	17.2	572	1	DPY2_MOUSE	DIHYDROPYRIMIDINASE RE	1.25e+02
124	5	17.2	382	1	LGT_MYCCE	PROBABLE PROLIPOPROTEI	1.25e+02	197	5	17.2	572	1	DPY1_RAT	DIHYDROPYRIMIDINASE RE	1.25e+02
125	5	17.2	382	1	AYRB_RAT	ACTIVIN RECEPTOR TYPE	1.25e+02	198	5	17.2	572	1	DPY1_HUMAN	DIHYDROPYRIMIDINASE RE	1.25e+02
126	5	17.2	383	1	LA_AEAL	LA PROTEIN HOMOLOG (LA	1.25e+02	199	5	17.2	572	1	DPY1_MOUSE	DIHYDROPYRIMIDINASE RE	1.25e+02
127	5	17.2	386	1	SYT5_RAT	SYNAPTOTAGMIN V.	1.25e+02	200	5	17.2	572	1	DPY2_CHICK	DIHYDROPYRIMIDINASE RE	1.25e+02
128	5	17.2	389	1	VMSA_HPBVM	MAJOR SURFACE ANTIGEN	1.25e+02	201	5	17.2	572	1	DPY2_RAT	DIHYDROPYRIMIDINASE RE	1.25e+02
129	5	17.2	392	1	PRGH_SALTY	PRGH PROTEIN.	1.25e+02	202	5	17.2	572	1	DPY2_HUMAN	DIHYDROPYRIMIDINASE RE	1.25e+02
130	5	17.2	394	1	MYC1_CYPCA	MYC I PROTEIN (C-MYC I	1.25e+02	203	5	17.2	575	1	SYP_ARCFU	PROLYL-TRNA SYNTHETASE	1.25e+02
131	5	17.2	399	1	MYC2_CARAU	MYC II PROTEIN (C-MYC)	1.25e+02	204	5	17.2	577	1	NRDD_BPT4	ANAEROBIC RIBONUCLEOSI	1.25e+02
132	5	17.2	401	1	MYC2_CYPCA	MYC II PROTEIN (C-MYC)	1.25e+02	205	5	17.2	577	1	PHBC_RHIME	POLY-BETA-HYDROXYBUTYR	1.25e+02
133	5	17.2	406	1	MYC_BRARE	MYC PROTEIN (C-MYC)	1.25e+02	206	5	17.2	578	1	YFAA_ECOLI	HYPOTHETICAL 64.5 KD P	1.25e+02
134	5	17.2	407	1	RL4_DROME	60S RIBOSOMAL PROTEIN	1.25e+02	207	5	17.2	583	1	CH12_CANAL	CHITINASE 2 PRECURSOR	1.25e+02
135	5	17.2	411	1	CAPA_BACAN	CAPA PROTEIN.	1.25e+02	208	5	17.2	588	1	DECA_DROME	DECAPENTAPLEGIC PROTEI	1.25e+02
136	5	17.2	417	1	HS47_RAT	47 KD HEAT SHOCK PROTE	1.25e+02	209	5	17.2	589	1	ESTR_CHICK	ESTROGEN RECEPTOR (ER)	1.25e+02
137	5	17.2	417	1	HS47_MOUSE	47 KD HEAT SHOCK PROTE	1.25e+02	210	5	17.2	593	1	DECA_DROSI	DECAPENTAPLEGIC PROTEI	1.25e+02
138	5	17.2	419	1	YHAP_BACSU	HYPOTHETICAL 45.4 KD P	1.25e+02	211	5	17.2	595	1	ESTR_PIG	ESTROGEN RECEPTOR (ER)	1.25e+02
139	5	17.2	419	1	MAT3_HUMAN	MATRIN 3 (FRAGMENT)	1.25e+02	212	5	17.2	597	1	YD13_ANTCR	DYNEIN INTERMEDIATE CH	1.25e+02
140	5	17.2	422	1	TEBL_BSPSP	TERMINASE LARGE SUBUNIT	1.25e+02	213	5	17.2	599	1	ESTR_MOUSE	ESTROGEN RECEPTOR (ER)	1.25e+02
141	5	17.2	424	1	I131_MOUSE	INTERLEUKIN-13 RECEPTO	1.25e+02	214	5	17.2	600	1	ESTR_RAT	ESTROGEN RECEPTOR (ER)	1.25e+02
142	5	17.2	424	1	IVD_RAT	ISOVALERYL-COA DEHYDRO	1.25e+02	215	5	17.2	603	1	GLMS_THERH	GLUCOSAMINE--FRUCTOSE-	1.25e+02
143	5	17.2	426	1	EXLP_TOBAC	PISTIL-SPECIFIC EXTENS	1.25e+02	216	5	17.2	605	1	YHC8_YEAST	HYPOTHETICAL 63.8 KD P	1.25e+02
144	5	17.2	428	1	FOLC_LACCA	POLYIPOLYGLUTAMATE SYN	1.25e+02	217	5	17.2	606	1	SP2_HUMAN	TRANSCRIPTION FACTOR S	1.25e+02
145	5	17.2	434	1	YBAR_BACSU	HYPOTHETICAL 46.4 KD P	1.25e+02	218	5	17.2	611	1	XYNA_PSEFL	ENDO-1,4-BETA-XYLANASE	1.25e+02
146	5	17.2	438	1	IE63_EBV	TRANSCRIPTIONAL REGULA	1.25e+02	219	5	17.2	612	1	OCTC_BOVIN	PEROXISOMAL CARNITINE	1.25e+02
147	5	17.2	444	1	NUPD_EMENI	NUCLEAR MIGRATION PROT	1.25e+02	220	5	17.2	614	1	EAR1_HUMAN	V-ERBA RELATED PROTEIN	1.25e+02
148	5	17.2	446	1	KCCL_YEAST	CALCIUM/CALMODULIN-DEP	1.25e+02	221	5	17.2	617	1	SG2_MOUSE	SECRETOGHRANIN II PRECU	1.25e+02
149	5	17.2	447	1	SLF1_YEAST	SLF1 PROTEIN.	1.25e+02	222	5	17.2	617	1	SG2_HUMAN	SECRETOGHRANIN II PRECU	1.25e+02
150	5	17.2	449	1	ALG1_YEAST	BETA-MANNOSYLTRANSFERA	1.25e+02	223	5	17.2	619	1	SG2_RAT	SECRETOGHRANIN II PRECU	1.25e+02
151	5	17.2	449	1	GUNA_CELFI	ENDOGLUCANASE A PRECUR	1.25e+02	224	5	17.2	625	1	AMYG_NEUCR	GLUCOAMYLASE PRECURSOR	1.25e+02
152	5	17.2	451	1	GAGY_DROME	RETROVIRUS-RELATED GAG	1.25e+02	225	5	17.2	634	1	HWP1_CANAL	HYPHAL WALL PROTEIN 1	1.25e+02
153	5	17.2	452	1	BMF5_MOUSE	BONE MORPHOGENETIC PRO	1.25e+02	226	5	17.2	636	1	YNR6_YEAST	HYPOTHETICAL 67.4 KD P	1.25e+02
154	5	17.2	454	1	BMF5_HUMAN	BONE MORPHOGENETIC PRO	1.25e+02	227	5	17.2	649	1	PUBL1_TORDE	HYPOPHOSPHOLIPASE PREC	1.25e+02
155	5	17.2	455	1	NMT_YEAST	GLYCYLPEPTIDE N-TETRAD	1.25e+02	228	5	17.2	650	1	DMR9_MOUSE	DMR-N9 PROTEIN.	1.25e+02
156	5	17.2	459	1	YGW9_YEAST	HYPOTHETICAL 52.0 KD P	1.25e+02	229	5	17.2	658	1	DNLJ_MYCPN	DNA LIGASE (EC 6.5.1.2	1.25e+02
157	5	17.2	460	1	TEBH_EUCEL	TELOMERE-BINDING PROTE	1.25e+02	230	5	17.2	658	1	PAK1_SCHPO	SERINE/THREONINE-PROTE	1.25e+02
158	5	17.2	467	1	UNB6_CAEEL	TRANSCRIPTION FACTOR U	1.25e+02	231	5	17.2	659	1	VHHS_YEAST	HYPOTHETICAL 75.9 KD P	1.25e+02
159	5	17.2	480	1	GLTD_AZOB	GLUTAMATE SYNTHASE [NA	1.25e+02	232	5	17.2	666	1	ZP2_FABIT	ZONA PELLUCIDA SERUM-B	1.25e+02
160	5	17.2	482	1	Y269_TREPA	HYPOTHETICAL PROTEIN T	1.25e+02	233	5	17.2	669	1	GSH1_SCHPO	GLUTAMATE--CYSTEINE LI	1.25e+02
161	5	17.2	485	1	IA12_LYCES	1-AMINOCYCLOPROPANE-1-	1.25e+02	234	5	17.2	670	1	KGPA_BOVIN	CGMP-DEPENDENT PROTEIN	1.25e+02
162	5	17.2	485	1	ATPB_CYAPA	ATP SYNTHASE BETA CHAI	1.25e+02	235	5	17.2	676	1	LSHR_CALJA	LUTROPIN-CHORIOGNADOT	1.25e+02
163	5	17.2	489	1	MSG5_YEAST	PROTEIN-TYROSINE PHOSP	1.25e+02	236	5	17.2	686	1	PLB3_YEAST	PUTATIVE LYPSOPHOSPHOL	1.25e+02
164	5	17.2	490	1	YMOE_BACSU	HYPOTHETICAL 54.0 KD P	1.25e+02	237	5	17.2	686	1	KGFB_HUMAN	CGMP-DEPENDENT PROTEIN	1.25e+02
165	5	17.2	493	1	YNL9_YEAST	HYPOTHETICAL 56.5 KD P	1.25e+02	238	5	17.2	686	1	KGFB_BOVIN	CGMP-DEPENDENT PROTEIN	1.25e+02
166	5	17.2	493	1	SR53_HORVU	SIGNAL RECOGNITION PAR	1.25e+02	239	5	17.2	696	1	LSHR_PIG	LUTROPIN-CHORIOGNADOT	1.25e+02
167	5	17.2	498	1	YDHF_SCHPO	HYPOTHETICAL 55.3 KD P	1.25e+02	240	5	17.2	699	1	SGS1_YEAST	SPORULATION PROTEIN GS	1.25e+02
168	5	17.2	501	1	VLI_PAPVE	MAJOR CAPSID PROTEIN L	1.25e+02	241	5	17.2	699	1	LSHR_HUMAN	LUTROPIN-CHORIOGNADOT	1.25e+02
169	5	17.2	503	1	PR19_YEAST	PRE-MRNA SPLICING FACT	1.25e+02	242	5	17.2	700	1	HS83_LEITN	HEAT SHOCK PROTEIN 83	1.25e+02



243	5	17.2	701	1	YG5L_YEAST	1.25e+02	1.25e+02	316	5	17.2	1022	1	ATNA_TORCA	SODIUM/POTASSIUM-TRANS	1.25e+02
244	5	17.2	701	1	LSHR_BOVIN	1.25e+02	1.25e+02	317	5	17.2	1030	1	FBP1_DROME	FAT-BODY PROTEIN-1 PRE	1.25e+02
245	5	17.2	701	1	HS83_LEIAM	1.25e+02	1.25e+02	318	5	17.2	1039	1	YAF3_SCHPO	HYPOHETICAL 118.6 KD	1.25e+02
246	5	17.2	705	1	LYTB_BACSU	1.25e+02	1.25e+02	319	5	17.2	1053	1	UBA3_WHEAT	UBIQUITIN-ACTIVATING E	1.25e+02
247	5	17.2	706	1	PLB2_YEAST	1.25e+02	1.25e+02	320	5	17.2	1120	1	AEFA_ECOLI	AEFA PROTEIN.	1.25e+02
248	5	17.2	708	1	PEY1_HUMAN	1.25e+02	1.25e+02	321	5	17.2	1125	1	Y079_HUMAN	HYPOHETICAL PROTEIN K	1.25e+02
249	5	17.2	716	1	HS83_DROAV	1.25e+02	1.25e+02	322	5	17.2	1131	1	YPO4_CAEEL	HYPOHETICAL 127.2 KD	1.25e+02
250	5	17.2	717	1	HS83_DROME	1.25e+02	1.25e+02	323	5	17.2	1132	1	YHC1_YEAST	NUT1 PROTEIN.	1.25e+02
251	5	17.2	726	1	PLOI_BOVIN	1.25e+02	1.25e+02	324	5	17.2	1146	1	YK82_YEAST	HYPOHETICAL 130.0 KD	1.25e+02
252	5	17.2	727	1	PLOI_HUMAN	1.25e+02	1.25e+02	325	5	17.2	1169	1	YK82_YEAST	HYPOHETICAL 122.2 KD	1.25e+02
253	5	17.2	732	1	ATN9_ECOLI	1.25e+02	1.25e+02	326	5	17.2	1169	1	RAD5_YEAST	DNA REPAIR PROTEIN RAD	1.25e+02
254	5	17.2	734	1	YD29_SCHPO	1.25e+02	1.25e+02	327	5	17.2	1177	1	JAK_DROME	TYROSINE-PROTEIN KINAS	1.25e+02
255	5	17.2	738	1	CC27_YEAST	1.25e+02	1.25e+02	328	5	17.2	1194	1	APAF_HUMAN	APOTOTIC PROTEASE ACT	1.25e+02
256	5	17.2	762	1	TRPG_NEUCR	1.25e+02	1.25e+02	329	5	17.2	1244	1	YPB3_CAEEL	HYPOHETICAL 139.1 KD	1.25e+02
257	5	17.2	776	1	ARCB_ECOLI	1.25e+02	1.25e+02	330	5	17.2	1260	1	ALS1_CANAL	AGGLUTININ-LIKE PROTEIN	1.25e+02
258	5	17.2	790	1	SMV2_YEAST	1.25e+02	1.25e+02	331	5	17.2	1296	1	BXG_CLOBO	BOTULININ-NEUROTOXIN T	1.25e+02
259	5	17.2	801	1	VP5_RDV	1.25e+02	1.25e+02	332	5	17.2	1301	1	DDX9_CAEEL	PROBABLE ATP-DEPENDENT	1.25e+02
260	5	17.2	810	1	PLMN_ERIEU	1.25e+02	1.25e+02	333	5	17.2	1336	1	SEC3_YEAST	EXOCYST COMPLEX COMPON	1.25e+02
261	5	17.2	820	1	TRBE_AGR16	1.25e+02	1.25e+02	334	5	17.2	1337	1	YDM5_SCHPO	HYPOHETICAL 150.8 KD	1.25e+02
262	5	17.2	822	1	GYRA_STRPN	1.25e+02	1.25e+02	335	5	17.2	1374	1	POLG_ECLIG	GENOME POLYPROTEIN [CO	1.25e+02
263	5	17.2	823	1	CC27_HUMAN	1.25e+02	1.25e+02	336	5	17.2	1379	1	MET_MOUSE	HEPATOCYTE GROWTH FACT	1.25e+02
264	5	17.2	827	1	MOF_DROME	1.25e+02	1.25e+02	337	5	17.2	1390	1	MET_HUMAN	HEPATOCYTE GROWTH FACT	1.25e+02
265	5	17.2	830	1	EF2_DICD1	1.25e+02	1.25e+02	338	5	17.2	1444	1	RRPL_RDV	RNA-DIRECTED RNA POLYM	1.25e+02
266	5	17.2	842	1	EF2_YEAST	1.25e+02	1.25e+02	339	5	17.2	1592	1	GF2_STRDO	GLUCOSYLTRANSFERASE-1	1.25e+02
267	5	17.2	842	1	PH8H_VICFA	1.25e+02	1.25e+02	340	5	17.2	1597	1	GF1_STRDO	GLUCOSYLTRANSFERASE-1	1.25e+02
268	5	17.2	846	1	ENV_HV2SB	1.25e+02	1.25e+02	341	5	17.2	1640	1	CUH2_HUMAN	CLATHRIN HEAVY CHAIN 2	1.25e+02
269	5	17.2	851	1	ENV_HV2G1	1.25e+02	1.25e+02	342	5	17.2	1675	1	CUH1_HUMAN	CLATHRIN HEAVY CHAIN 1	1.25e+02
270	5	17.2	851	1	ENV_HV2D1	1.25e+02	1.25e+02	343	5	17.2	1675	1	CUH1_BOVIN	CLATHRIN HEAVY CHAIN 1	1.25e+02
271	5	17.2	854	1	XR52_YEAST	1.25e+02	1.25e+02	344	5	17.2	1675	1	CUH_RAT	CLATHRIN HEAVY CHAIN.	1.25e+02
272	5	17.2	854	1	ENV_SIVAI	1.25e+02	1.25e+02	345	5	17.2	1790	1	USO1_YEAST	INTRACELLULAR PROTEIN	1.25e+02
273	5	17.2	855	1	NOLI_HUMAN	1.25e+02	1.25e+02	346	5	17.2	1805	1	NES1_RAT	NESTIN.	1.25e+02
274	5	17.2	856	1	ENV_HV2NZ	1.25e+02	1.25e+02	347	5	17.2	1848	1	CBPA_CLOCL	CELLULOSE BINDING PROT	1.25e+02
275	5	17.2	857	1	ENV_HV2KR	1.25e+02	1.25e+02	348	5	17.2	1850	1	VIT2_CHICK	VITELLOGENIN II PRECUR	1.25e+02
276	5	17.2	858	1	ENV_HV2RO	1.25e+02	1.25e+02	349	5	17.2	1857	1	MYSN_HUMAN	MYOSIN HEAVY CHAIN, SM	1.25e+02
277	5	17.2	859	1	ENV_HV2CA	1.25e+02	1.25e+02	350	5	17.2	1959	1	MYSN_CHICK	MYOSIN HEAVY CHAIN, NO	1.25e+02
278	5	17.2	859	1	ENV_HV2ST	1.25e+02	1.25e+02	351	5	17.2	1960	1	MYSN_HUMAN	MYOSIN HEAVY CHAIN, NO	1.25e+02
279	5	17.2	860	1	KDPD_MCTU	1.25e+02	1.25e+02	352	5	17.2	1972	1	MYST_RABIT	MYOSIN HEAVY CHAIN, SM	1.25e+02
280	5	17.2	860	1	ENV_HV2BE	1.25e+02	1.25e+02	353	5	17.2	1976	1	MYSO_HUMAN	MYOSIN HEAVY CHAIN, NO	1.25e+02
281	5	17.2	863	1	YK44_YEAST	1.25e+02	1.25e+02	354	5	17.2	1978	1	MYSG_CHICK	MYOSIN HEAVY CHAIN, GI	1.25e+02
282	5	17.2	863	1	GYRA_CAMJE	1.25e+02	1.25e+02	355	5	17.2	2017	1	MYSN_DROME	MYOSIN HEAVY CHAIN, NO	1.25e+02
283	5	17.2	864	1	LOX2_PEA	1.25e+02	1.25e+02	356	5	17.2	2026	1	CYAA_YEAST	ADENYLATE CYCLASE (EC	1.25e+02
284	5	17.2	864	1	KLFK_HUMAN	1.25e+02	1.25e+02	357	5	17.2	2116	1	RRPL_SYNV	RNA POLYMERASE BETA SU	1.25e+02
285	5	17.2	864	1	FAS_YEAST	1.25e+02	1.25e+02	358	5	17.2	2148	1	VIT1_AEDAE	VITELLOGENIN A1 PRECUR	1.25e+02
286	5	17.2	875	1	IMB1_RAT	1.25e+02	1.25e+02	359	5	17.2	2149	1	RRPL_RVFPVZ	RNA-DIRECTED RNA POLYM	1.25e+02
287	5	17.2	876	1	IMB1_MOUSE	1.25e+02	1.25e+02	360	5	17.2	2150	1	POLG_HRVZ	GENOME POLYPROTEIN [CO	1.25e+02
288	5	17.2	876	1	IMB1_HUMAN	1.25e+02	1.25e+02	361	5	17.2	2182	1	POLG_COXB1	GENOME POLYPROTEIN [CO	1.25e+02
289	5	17.2	879	1	MAB1_BOVIN	1.25e+02	1.25e+02	362	5	17.2	2183	1	POLG_COXB4	GENOME POLYPROTEIN [CO	1.25e+02
290	5	17.2	881	1	NIA1_PRAVU	1.25e+02	1.25e+02	363	5	17.2	2185	1	POLG_COXB5	GENOME POLYPROTEIN [CO	1.25e+02
291	5	17.2	886	1	NIA1_SOYBN	1.25e+02	1.25e+02	364	5	17.2	2185	1	POLG_SVDVH	GENOME POLYPROTEIN [CO	1.25e+02
292	5	17.2	886	1	ODP1_ECOLI	1.25e+02	1.25e+02	365	5	17.2	2185	1	POLG_SVDVH	GENOME POLYPROTEIN [CO	1.25e+02
293	5	17.2	887	1	ACOC_CAEEL	1.25e+02	1.25e+02	366	5	17.2	2185	1	POLG_COXB3	GENOME POLYPROTEIN [CO	1.25e+02
294	5	17.2	890	1	NIA2_PRAVU	1.25e+02	1.25e+02	367	5	17.2	2194	1	POLG_HE701	GENOME POLYPROTEIN [CO	1.25e+02
295	5	17.2	890	1	NIA2_SOYBN	1.25e+02	1.25e+02	368	5	17.2	2201	1	POLG_CXA9	GENOME POLYPROTEIN [CO	1.25e+02
296	5	17.2	893	1	KNRB_YEAST	1.25e+02	1.25e+02	369	5	17.2	2231	1	SEN1_YEAST	TRNA-SPLICING ENDONUCL	1.25e+02
297	5	17.2	898	1	NIA1_BETVE	1.25e+02	1.25e+02	370	5	17.2	2284	1	POLL1_GFLV	RNA1 POLYPROTEIN (253	1.25e+02
298	5	17.2	900	1	NIA1_LOTJA	1.25e+02	1.25e+02	371	5	17.2	2325	1	PGG2_RAT	CHONDROITIN SULFATE PR	1.25e+02
299	5	17.2	900	1	FOX2_YEAST	1.25e+02	1.25e+02	372	5	17.2	2332	1	POLG_FMDVO	GENOME POLYPROTEIN [CO	1.25e+02
300	5	17.2	904	1	NIA1_TOBAC	1.25e+02	1.25e+02	373	5	17.2	2332	1	POLG_FMDVA	GENOME POLYPROTEIN [CO	1.25e+02
301	5	17.2	904	1	NIA2_TOBAC	1.25e+02	1.25e+02	374	5	17.2	2333	1	POLG_FMDV1	GENOME POLYPROTEIN [CO	1.25e+02
302	5	17.2	909	1	NIA1_PETHY	1.25e+02	1.25e+02	375	5	17.2	2336	1	POLG_FMDVZ	GENOME POLYPROTEIN [CO	1.25e+02
303	5	17.2	911	1	NIA1_LYCES	1.25e+02	1.25e+02	376	5	17.2	2339	1	RFC1_PLAFA	DNA-DIRECTED RNA POLYM	1.25e+02
304	5	17.2	911	1	NIA1_BRANA	1.25e+02	1.25e+02	377	5	17.2	2493	1	YBA4_YEAST	HYPOHETICAL 287.5 KD	1.25e+02
305	5	17.2	911	1	NIA2_BRANA	1.25e+02	1.25e+02	378	5	17.2	2515	1	TUD_DROME	INOSITOL 1,4,5-TRISPHO	1.25e+02
306	5	17.2	917	1	NIA1_ARATH	1.25e+02	1.25e+02	379	5	17.2	2833	1	IP3R_DROME	MATERNAL TUDOR RNA	1.25e+02
307	5	17.2	918	1	NIA1_CUCMA	1.25e+02	1.25e+02	380	5	17.2	2875	1	RRPL_TSWV1	RNA-DIRECTED RNA POLYM	1.25e+02
308	5	17.2	920	1	NIA1_CICIN	1.25e+02	1.25e+02	381	5	17.2	2960	1	AHNK_HUMAN	NEUROBLAST DIFFERENTIA	1.25e+02
309	5	17.2	923	1	YOL3_CAEEL	1.25e+02	1.25e+02	382	5	17.2	3066	1	POLG_SBMVG	GENOME POLYPROTEIN [CO	1.25e+02
310	5	17.2	926	1	NIA1_SPTOL	1.25e+02	1.25e+02	383	5	17.2	3591	1	FNAB_BORPE	FILAMENTOUS HEMAGGLUTI	1.25e+02
311	5	17.2	928	1	NRP_XENLA	1.25e+02	1.25e+02	384	5	17.2	4544	1	LRP1_HUMAN	LOW-DENSITY LIPOPROTEI	1.25e+02
312	5	17.2	960	1	CAP3_SORVU	1.25e+02	1.25e+02	385	4	13.8	33	1	VC12_CHLRE	HYPOHETICAL 3.3 KD PR	2.87e+03
313	5	17.2	1011	1	POLG_FMDVT	1.25e+02	1.25e+02	386	5	17.2	42	1	VP10_BPPH6	P10 PROTEIN.	2.87e+03
314	5	17.2	1016	1	POLG_MNV2	1.25e+02	1.25e+02	387	4	13.8	49	1	YDCB_ECOLI	HYPOHETICAL 5.6 KD PR	2.87e+03
315	5	17.2	1018	1	FNBA_STAAU	1.25e+02	1.25e+02	388	4	13.8	55	1	VSE2_TRYBR	VARIANT SURFACE GLYCOP	2.87e+03

389	4	13.8	58	1	YBAS_PHAVU	HYPOTHETICAL BASIC POL	2.87e+03	462	4	13.8	168	1	YNXA_BACSU	HYPOTHETICAL PROTEIN I	2.87e+03
390	4	13.8	59	1	YORY_LISMO	HYPOTHETICAL 6.9 KD PR	2.87e+03	463	4	13.8	168	1	YE26_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
391	4	13.8	65	1	VSH_HRSV1	SMALL HYDROPHOBIC PROT	2.87e+03	464	4	13.8	169	1	YCF3_PINTH	HYPOTHETICAL 19.5 KD P	2.87e+03
392	4	13.8	68	1	Y03B_BPT4	HYPOTHETICAL 8.2 KD PR	2.87e+03	465	4	13.8	170	1	YME1_THIFE	HYPOTHETICAL 18.9 KD P	2.87e+03
393	4	13.8	68	1	YOR3_FIVSD	HYPOTHETICAL 7.7 KD PR	2.87e+03	466	4	13.8	172	1	YNE6_YEAST	HYPOTHETICAL 19.4 KD P	2.87e+03
394	4	13.8	70	1	YORA_TTV1	HYPOTHETICAL 8.1 KD PR	2.87e+03	467	4	13.8	178	1	Y01J5_CABEL	HYPOTHETICAL CALCIUM-B	2.87e+03
395	4	13.8	70	1	YNCH_ECOLI	HYPOTHETICAL 8.3 KD PR	2.87e+03	468	4	13.8	178	1	Y0GE_MYCTU	HYPOTHETICAL 19.5 KD P	2.87e+03
396	4	13.8	71	1	Y0JD_ECOLI	HYPOTHETICAL 7.9 KD PR	2.87e+03	469	4	13.8	179	1	Y0IB_ECOLI	HYPOTHETICAL 20.8 KD P	2.87e+03
397	4	13.8	72	1	YPTA_INVA	HYPOTHETICAL P7A PROTE	2.87e+03	470	4	13.8	180	1	YB0A_SCHPO	HYPOTHETICAL 20.9 KD P	2.87e+03
398	4	13.8	78	1	Y43_BPR51	HYPOTHETICAL PROTEIN O	2.87e+03	471	4	13.8	181	1	Y041_NPVAC	HYPOTHETICAL 21.1 KD P	2.87e+03
399	4	13.8	82	1	YVCC_VACCC	HYPOTHETICAL 9.2 KD PR	2.87e+03	472	4	13.8	181	1	YKN9_YEAST	VERY HYPOTHETICAL 19.7	2.87e+03
400	4	13.8	84	1	YM66_CAEEL	HYPOTHETICAL 8.8 KD PR	2.87e+03	473	4	13.8	183	1	YEBR_ECOLI	HYPOTHETICAL 20.3 KD P	2.87e+03
401	4	13.8	84	1	YN77_YEAST	HYPOTHETICAL 9.7 KD PR	2.87e+03	474	4	13.8	184	1	YDB5_SCHPO	HYPOTHETICAL 22.4 KD P	2.87e+03
402	4	13.8	87	1	Y14B_BPT4	HYPOTHETICAL 10.1 KD P	2.87e+03	475	4	13.8	184	1	Y080_ACTPL	HYPOTHETICAL 21.6 KD P	2.87e+03
403	4	13.8	87	1	YAL2_BACLI	HYPOTHETICAL 9.5 KD PR	2.87e+03	476	4	13.8	185	1	YHBN_ECOLI	17.3 KD PROTEIN IN MUR	2.87e+03
404	4	13.8	91	1	YBEM_ECOLI	HYPOTHETICAL 10.4 KD P	2.87e+03	477	4	13.8	185	1	YEM4_YEAST	HYPOTHETICAL 21.2 KD P	2.87e+03
405	4	13.8	91	1	YBEM_ECOLI	HYPOTHETICAL 10.5 KD P	2.87e+03	478	4	13.8	186	1	YALD_TRYBB	HYPOTHETICAL 21.2 KD PRO	2.87e+03
406	4	13.8	95	1	Y129_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	479	4	13.8	186	1	YIHI_HAEIN	HYPOTHETICAL PROTEIN B	2.87e+03
407	4	13.8	99	1	YIIS_ECOLI	HYPOTHETICAL 10.8 KD P	2.87e+03	480	4	13.8	190	1	Y001_BORBU	HYPOTHETICAL PROTEIN H	2.87e+03
408	4	13.8	99	1	Y048_GUITH	HYPOTHETICAL 11.6 KD P	2.87e+03	481	4	13.8	190	1	YRDC_ECOLI	HYPOTHETICAL 20.8 KD P	2.87e+03
409	4	13.8	103	1	YPC6_CLAPU	HYPOTHETICAL 11.4 KD P	2.87e+03	482	4	13.8	191	1	YRAP_ECOLI	HYPOTHETICAL 20.0 KD P	2.87e+03
410	4	13.8	105	1	YR7G_ECOLI	HYPOTHETICAL 12.0 KD P	2.87e+03	483	4	13.8	191	1	YAF8_CABEL	HYPOTHETICAL 22.6 KD P	2.87e+03
411	4	13.8	107	1	YXQ2_PORPU	HYPOTHETICAL 12.3 KD P	2.87e+03	484	4	13.8	196	1	YAGK_ECOLI	HYPOTHETICAL 23.3 KD P	2.87e+03
412	4	13.8	107	1	Y146_SYNY3	HYPOTHETICAL 11.9 KD P	2.87e+03	485	4	13.8	196	1	YBK7_YEAST	HYPOTHETICAL 23.0 KD P	2.87e+03
413	4	13.8	107	1	Y146_SYNY3	HYPOTHETICAL 12.1 KD P	2.87e+03	486	4	13.8	197	1	YGE7_YEAST	HYPOTHETICAL 22.8 KD P	2.87e+03
414	4	13.8	107	1	Y146_SYNY3	VERY HYPOTHETICAL 11.8	2.87e+03	487	4	13.8	199	1	YP24_RTBPV	HYPOTHETICAL 24.8 KD PRO	2.87e+03
415	4	13.8	108	1	YG16_YEAST	HYPOTHETICAL 12.4 KD P	2.87e+03	488	4	13.8	201	1	Y146_NPVAC	HYPOTHETICAL 22.9 KD P	2.87e+03
416	4	13.8	108	1	YB00_YEAST	HYPOTHETICAL 12.8 KD P	2.87e+03	489	4	13.8	201	1	YJAA_YEAST	HYPOTHETICAL 23.7 KD P	2.87e+03
417	4	13.8	109	1	YH37_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	490	4	13.8	203	1	VP10_BPRD	PROTEIN P10.	2.87e+03
418	4	13.8	110	1	Y152_STRCO	INSERTION ELEMENT IS11	2.87e+03	491	4	13.8	204	1	YPOL_ACICA	HYPOTHETICAL PROTEIN I	2.87e+03
419	4	13.8	110	1	Y0JH_ECOLI	HYPOTHETICAL 12.3 KD P	2.87e+03	492	4	13.8	206	1	YPKP_BACSU	HYPOTHETICAL 23.3 KD P	2.87e+03
420	4	13.8	111	1	Y0C8_YEAST	HYPOTHETICAL 12.9 KD P	2.87e+03	493	4	13.8	208	1	XLR_MOUSE	X-LINKED LYMPHOCYTE-RE	2.87e+03
421	4	13.8	115	1	YBR2_YEAST	HYPOTHETICAL 13.2 KD P	2.87e+03	494	4	13.8	210	1	YPS1_YEAST	GTP-BINDING PROTEIN YP	2.87e+03
422	4	13.8	116	1	Y392_YEAST	HYPOTHETICAL 13.4 KD P	2.87e+03	495	4	13.8	211	1	Y121_BURCE	INSERTION ELEMENT IS40	2.87e+03
423	4	13.8	117	1	YIAB_ECOLI	HYPOTHETICAL 13.0 KD P	2.87e+03	496	4	13.8	212	1	YCDC_ECOLI	HYPOTHETICAL TRANScrip	2.87e+03
424	4	13.8	117	1	YNGK_CAEEL	HYPOTHETICAL 12.9 KD P	2.87e+03	497	4	13.8	212	1	YCD4_YEAST	HYPOTHETICAL 23.6 KD P	2.87e+03
425	4	13.8	118	1	Y151_ECOLI	HYPOTHETICAL 13.3 KD P	2.87e+03	498	4	13.8	212	1	VS10_R0THC	MINOR OUTER CAPSID PRO	2.87e+03
426	4	13.8	118	1	Y035_BPLIH	INSERTION ELEMENT IS5	2.87e+03	499	4	13.8	217	1	YB47_YEAST	HYPOTHETICAL 24.2 KD P	2.87e+03
427	4	13.8	119	1	YRAN_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	500	4	13.8	218	1	Y381_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03
428	4	13.8	121	1	Y162_BURCE	INSERTION ELEMENT IS40	2.87e+03	501	4	13.8	220	1	YNLO_YEAST	HYPOTHETICAL 25.4 KD P	2.87e+03
429	4	13.8	125	1	Y055_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03	502	4	13.8	220	1	YCC7_ECOLI	HYPOTHETICAL 24.6 KD P	2.87e+03
430	4	13.8	126	1	YAM2_YEAST	VERY HYPOTHETICAL 13.2	2.87e+03	503	4	13.8	220	1	YICG_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03
431	4	13.8	130	1	Y042_YEAST	HYPOTHETICAL 14.9 KD P	2.87e+03	504	4	13.8	222	1	Y434_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
432	4	13.8	130	1	Y060_BPT4	HYPOTHETICAL 15.1 KD P	2.87e+03	505	4	13.8	222	1	XYLJ_PSEPU	2-HYDROXYPENT-2,4-DIEN	2.87e+03
433	4	13.8	131	1	YV33C_CAPVK	T3C PROTEIN.	2.87e+03	506	4	13.8	222	1	YEC5_ECOLI	HYPOTHETICAL AMINO-ACI	2.87e+03
434	4	13.8	131	1	YVCU_BACSU	HYPOTHETICAL 15.3 KD P	2.87e+03	507	4	13.8	224	1	YE9A_SCHPO	HYPOTHETICAL 24.1 KD P	2.87e+03
435	4	13.8	131	1	YUG4_YEAST	HYPOTHETICAL 13.9 KD P	2.87e+03	508	4	13.8	224	1	Y091_NPVAC	HYPOTHETICAL 24.1 KD P	2.87e+03
436	4	13.8	132	1	YNO3_PARDE	HYPOTHETICAL 14.7 KD P	2.87e+03	509	4	13.8	225	1	YMI3_CAEEL	HYPOTHETICAL 25.2 KD P	2.87e+03
437	4	13.8	134	1	YI1Y_SALTY	HYPOTHETICAL 15.6 KD P	2.87e+03	510	4	13.8	225	1	YEAL_YEAST	HYPOTHETICAL 25.0 KD P	2.87e+03
438	4	13.8	138	1	YF68_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	511	4	13.8	227	1	Y995_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
439	4	13.8	140	1	Y01C_BACSU	HYPOTHETICAL 15.7 KD P	2.87e+03	512	4	13.8	229	1	Y014_ARCFU	HYPOTHETICAL PROTEIN A	2.87e+03
440	4	13.8	141	1	YNTR_ECOLI	ZN(II)-RESPONSIVE REGU	2.87e+03	513	4	13.8	229	1	YAG7_ECOLI	HYPOTHETICAL 24.3 KD P	2.87e+03
441	4	13.8	143	1	Y111_ADE41	HYPOTHETICAL 16.5 KD P	2.87e+03	514	4	13.8	229	1	Y4MB_RHISN	HYPOTHETICAL OUTER-MEM	2.87e+03
442	4	13.8	144	1	YG34_BPMV4	HYPOTHETICAL 15.6 KD P	2.87e+03	515	4	13.8	232	1	YDFE_SCHPO	HYPOTHETICAL 26.1 KD P	2.87e+03
443	4	13.8	145	1	VP3_BPCHP	PROTEIN VP3 (ORF3).	2.87e+03	516	4	13.8	232	1	YALV_TRYBB	HYPOTHETICAL 25.6 KD P	2.87e+03
444	4	13.8	147	1	YKDM_BACSU	PHAGE-LIKE ELEMENT PBS	2.87e+03	517	4	13.8	233	1	YJ08_YEAST	HYPOTHETICAL 26.9 KD P	2.87e+03
445	4	13.8	148	1	YEN4_YEAST	HYPOTHETICAL 17.1 KD P	2.87e+03	518	4	13.8	234	1	YCK1_BACSU	PROBABLE AMINO-ACID AB	2.87e+03
446	4	13.8	148	1	YGEK_ECOLI	HYPOTHETICAL TRANScrip	2.87e+03	519	4	13.8	234	1	YELB_SCHPO	POTATIVE RAS-RELATED P	2.87e+03
447	4	13.8	151	1	YBBJ_ECOLI	HYPOTHETICAL 16.8 KD P	2.87e+03	520	4	13.8	234	1	YPS2_YEAST	GTP-BINDING PROTEIN YP	2.87e+03
448	4	13.8	151	1	YFG3_YEAST	HYPOTHETICAL 17.5 KD P	2.87e+03	521	4	13.8	235	1	YJBM_ECOLI	HYPOTHETICAL 26.7 KD P	2.87e+03
449	4	13.8	152	1	YK71_YEAST	HYPOTHETICAL 17.0 KD P	2.87e+03	522	4	13.8	235	1	YKDP_BACSU	PHAGE-LIKE ELEMENT PBS	2.87e+03
450	4	13.8	152	1	YCB2_YEAST	HYPOTHETICAL 17.0 KD P	2.87e+03	523	4	13.8	238	1	YIDP_ECOLI	HYPOTHETICAL TRANScrip	2.87e+03
451	4	13.8	153	1	YORZ_LISMO	HYPOTHETICAL 16.9 KD L	2.87e+03	524	4	13.8	238	1	Y647_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03
452	4	13.8	153	1	YH21_YEAST	HYPOTHETICAL 17.2 KD P	2.87e+03	525	4	13.8	238	1	Y4TF_RHISN	PROBABLE AMINO-ACID AB	2.87e+03
453	4	13.8	153	1	YH17_YEAST	HYPOTHETICAL 17.1 KD P	2.87e+03	526	4	13.8	242	1	YURK_BACSU	HYPOTHETICAL TRANScrip	2.87e+03
454	4	13.8	157	1	YR37_MYCTU	HYPOTHETICAL 16.9 KD P	2.87e+03	527	4	13.8	243	1	YC23_CVAPA	HYPOTHETICAL 26.6 KD P	2.87e+03
455	4	13.8	158	1	YNI9_YEAST	HYPOTHETICAL 17.1 KD P	2.87e+03	528	4	13.8	244	1	YPHF_BACSU	HYPOTHETICAL 27.7 KD P	2.87e+03
456	4	13.8	158	1	YKPS_KLULA	HYPOTHETICAL KILLER PL	2.87e+03	529	4	13.8	245	1	YABS_BACSU	HYPOTHETICAL 26.8 KD P	2.87e+03
457	4	13.8	162	1	YFCT_ECOLI	HYPOTHETICAL 18.9 KD P	2.87e+03	530	4	13.8	245	1	YR42_CAEEL	PROBABLE G PROTEIN-CON	2.87e+03
458	4	13.8	163	1	YNZ5_CAEEL	HYPOTHETICAL 19.4 KD P	2.87e+03	531	4	13.8	245	1	Y355_HAEIN	HYPOTHETICAL ABC TRANS	2.87e+03
459	4	13.8	163	1	YAS5_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	532	4	13.8	248	1	WCAC_ECOLI	POTATIVE COLANIC ACID	2.87e+03
460	4	13.8	164	1	YF18_FOMP1	HYPOTHETICAL 18.0 KD P	2.87e+03	533	4	13.8	250	1	YFCS_ECOLI	HYPOTHETICAL FIMBRIAL	2.87e+03
461	4	13.8	166	1	Y18K_SSV1	HYPOTHETICAL 18.0 KD P	2.87e+03	534	4	13.8	250	1	VNST_PTPV	NONSTRUCTURAL PROTEIN	2.87e+03

535	4	13.8	253	1	YHCE_BACSU	HYPOTHETICAL 29.5 KD P	2.87e+03	608	4	13.8	335	1	Y094_CAEEL	HYPOTHETICAL 40.1 KD P	2.87e+03
536	4	13.8	256	1	YG2N_YEAST	HYPOTHETICAL 28.3 KD P	2.87e+03	609	4	13.8	335	1	XYNB_STRLL	ENDO-1,4-BETA-XYLANASE	2.87e+03
537	4	13.8	256	1	YRDA_ECOLI	HYPOTHETICAL 28.4 KD P	2.87e+03	610	4	13.8	336	1	YB56_XENLA	CYTOPLASMIC RNA-BINDING	2.87e+03
538	4	13.8	257	1	Y08C_MYCTU	HYPOTHETICAL 26.4 KD P	2.87e+03	611	4	13.8	338	1	YORT_TTV1	HYPOTHETICAL 38.7 KD P	2.87e+03
539	4	13.8	257	1	YB06_YEAST	HYPOTHETICAL 29.1 KD P	2.87e+03	612	4	13.8	341	1	YEYE_ECOLI	HYPOTHETICAL ABC TRANS	2.87e+03
540	4	13.8	257	1	YFJB_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	613	4	13.8	341	1	YH10_YEAST	HYPOTHETICAL 37.9 KD P	2.87e+03
541	4	13.8	265	1	YPG1_ZYMMO	HYPOTHETICAL 28.1 KD P	2.87e+03	614	4	13.8	342	1	YEAM_YEAST	HYPOTHETICAL 39.3 KD P	2.87e+03
542	4	13.8	265	1	YOR3_LELV	HYPOTHETICAL 30.6 KD P	2.87e+03	615	4	13.8	343	1	YJCR_ECOLI	HYPOTHETICAL 36.9 KD P	2.87e+03
543	4	13.8	267	1	YCHA_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	616	4	13.8	344	1	Y09A_MYCTU	HYPOTHETICAL 35.1 KD P	2.87e+03
544	4	13.8	268	1	YBGF_PSEPU	HYPOTHETICAL 28.7 KD P	2.87e+03	617	4	13.8	345	1	Y456_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03
545	4	13.8	268	1	YHC6_YEAST	HYPOTHETICAL 30.7 KD P	2.87e+03	618	4	13.8	346	1	YG2Q_YEAST	HYPOTHETICAL 38.8 KD P	2.87e+03
546	4	13.8	269	1	Y085_HUMAN	HYPOTHETICAL PROTEIN K	2.87e+03	619	4	13.8	346	1	YIC3_YEAST	HYPOTHETICAL 37.4 KD P	2.87e+03
547	4	13.8	271	1	YDIB_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	620	4	13.8	347	1	YCR6_ERWHE	HYPOTHETICAL 37.2 KD P	2.87e+03
548	4	13.8	271	1	YPOQ_KLEPN	HYPOTHETICAL PROTEIN I	2.87e+03	621	4	13.8	348	1	Y09U_MYCTU	HYPOTHETICAL 38.1 KD P	2.87e+03
549	4	13.8	271	1	YDUO_ECOLI	HYPOTHETICAL 30.7 KD P	2.87e+03	622	4	13.8	351	1	Y876_METJA	PUTATIVE ABC TRANSPORT	2.87e+03
550	4	13.8	274	1	Y374_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	623	4	13.8	351	1	Y246_MYCLE	HYPOTHETICAL 35.4 KD P	2.87e+03
551	4	13.8	274	1	YDM7_SCHPO	HYPOTHETICAL 30.8 KD P	2.87e+03	624	4	13.8	351	1	Y4VJ_RHISN	HYPOTHETICAL 39.2 KD P	2.87e+03
552	4	13.8	275	1	Y01M_MYCTU	HYPOTHETICAL 30.8 KD P	2.87e+03	625	4	13.8	357	1	YJUU_ECOLI	HYPOTHETICAL 39.8 KD P	2.87e+03
553	4	13.8	276	1	YG65_YEAST	HYPOTHETICAL 31.3 KD H	2.87e+03	626	4	13.8	358	1	YJGQ_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03
554	4	13.8	277	1	XKDB_BACSU	PHAGE-LIKE ELEMENT PBS	2.87e+03	627	4	13.8	359	1	YH99_YEAST	HYPOTHETICAL 39.4 KD P	2.87e+03
555	4	13.8	277	1	YIBQ_ECOLI	HYPOTHETICAL 30.7 KD P	2.87e+03	628	4	13.8	360	1	Y0FA_CAEEL	HYPOTHETICAL 39.7 KD P	2.87e+03
556	4	13.8	279	1	Y091_NPVOP	HYPOTHETICAL 29.3 KD P	2.87e+03	629	4	13.8	362	1	YCHF_HAEIN	PROBABLE GTP-BINDING P	2.87e+03
557	4	13.8	279	1	YGA5_YEAST	HYPOTHETICAL 31.8 KD P	2.87e+03	630	4	13.8	362	1	Y024_MYCPN	PROBABLE GTP-BINDING P	2.87e+03
558	4	13.8	280	1	Y103_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	631	4	13.8	364	1	YBR5_YEAST	HYPOTHETICAL 40.9 KD P	2.87e+03
559	4	13.8	281	1	Y246_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	632	4	13.8	364	1	YBWL_CAEEL	HYPOTHETICAL 41.5 KD P	2.87e+03
560	4	13.8	282	1	YEA5_YEAST	HYPOTHETICAL 31.4 KD P	2.87e+03	633	4	13.8	365	1	YJCI_YEAST	HYPOTHETICAL 41.2 KD P	2.87e+03
561	4	13.8	282	1	YKGG_ECOLI	HYPOTHETICAL 31.1 KD P	2.87e+03	634	4	13.8	365	1	YQ23_CAEEL	HYPOTHETICAL ZINC FING	2.87e+03
562	4	13.8	287	1	Y193_HAEIN	PUTATIVE ESTERASE/LIPA	2.87e+03	635	4	13.8	365	1	Y971_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
563	4	13.8	289	1	YXNB_BACSU	HYPOTHETICAL 31.8 KD P	2.87e+03	636	4	13.8	367	1	Y024_MYCGE	PROBABLE GTP-BINDING P	2.87e+03
564	4	13.8	290	1	Y43A_MYCPN	HYPOTHETICAL LIPOPROTE	2.87e+03	637	4	13.8	367	1	Y797_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
565	4	13.8	290	1	Y084_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	638	4	13.8	368	1	YGYI_YEAST	HYPOTHETICAL 41.0 KD P	2.87e+03
566	4	13.8	290	1	Y066_YEAST	HYPOTHETICAL 33.6 KD P	2.87e+03	639	4	13.8	371	1	YPC4_CAEEL	HYPOTHETICAL 41.9 KD P	2.87e+03
567	4	13.8	290	1	YAS7_METJA	PUTATIVE GLYCOSYL TRAN	2.87e+03	640	4	13.8	373	1	YCFD_ECOLI	HYPOTHETICAL 42.6 KD P	2.87e+03
568	4	13.8	291	1	YNEB_ECOLI	HYPOTHETICAL 31.9 KD P	2.87e+03	641	4	13.8	374	1	Y197_MYCGE	HYPOTHETICAL 43.9 KD P	2.87e+03
569	4	13.8	300	1	YADD_ECOLI	HYPOTHETICAL 34.6 KD P	2.87e+03	642	4	13.8	375	1	Y147_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03
570	4	13.8	304	1	Y180_MYCGE	HYPOTHETICAL ABC TRANS	2.87e+03	643	4	13.8	376	1	YB11_YEAST	HYPOTHETICAL 44.4 KD P	2.87e+03
571	4	13.8	305	1	YHCB_BACSU	HYPOTHETICAL ABC TRANS	2.87e+03	644	4	13.8	376	1	YG3W_YEAST	HYPOTHETICAL 44.3 KD P	2.87e+03
572	4	13.8	307	1	Y07L_MYCTU	HYPOTHETICAL 33.9 KD P	2.87e+03	645	4	13.8	377	1	Y4WD_RHISN	HYPOTHETICAL TRANSPORT	2.87e+03
573	4	13.8	307	1	YMC1_YEAST	MITOCHONDRIAL CARRIER	2.87e+03	646	4	13.8	377	1	Y147_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03
574	4	13.8	307	1	Y678_TREPA	HYPOTHETICAL PROTEIN T	2.87e+03	647	4	13.8	377	1	YMX4_CAEEL	HYPOTHETICAL 43.2 KD P	2.87e+03
575	4	13.8	308	1	YBBS_ECOLI	HYPOTHETICAL TRANSCRIPT	2.87e+03	648	4	13.8	382	1	YJ07_YEAST	HYPOTHETICAL 44.9 KD P	2.87e+03
576	4	13.8	308	1	Y880_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	649	4	13.8	382	1	Y077_METJA	HYPOTHETICAL PROTEIN H	2.87e+03
577	4	13.8	309	1	YOH1_AGOV1	HYPOTHETICAL 33.2 KD P	2.87e+03	650	4	13.8	382	1	Y753_HAEIN	HYPOTHETICAL PROTEIN M	2.87e+03
578	4	13.8	310	1	YKJ1_YEAST	36.1 KD PROTEIN IN BUD	2.87e+03	651	4	13.8	383	1	Y0X1_YEAST	HOMEOBOX PROTEIN YOX1	2.87e+03
579	4	13.8	310	1	Y4RE_RHISN	PUTATIVE INTEGRASE/REC	2.87e+03	652	4	13.8	385	1	YX1Y_YEAST	HOMEOBOX PROTEIN YOX1	2.87e+03
580	4	13.8	310	1	YBAS_ECOLI	HYPOTHETICAL 32.9 KD P	2.87e+03	653	4	13.8	385	1	YEHV_ECOLI	HYPOTHETICAL ABC TRANS	2.87e+03
581	4	13.8	310	1	YANB_ECOLI	HYPOTHETICAL TRANSCRIPT	2.87e+03	654	4	13.8	385	1	YCT2_BACFI	HYPOTHETICAL 43.4 KD P	2.87e+03
582	4	13.8	313	1	YCHB_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03	655	4	13.8	385	1	YMD2_ECOLI	HYPOTHETICAL 44.7 KD P	2.87e+03
583	4	13.8	314	1	Y4TP_RHISN	PROBABLE PEPTIDE ABC T	2.87e+03	656	4	13.8	387	1	YJFC_ECOLI	HYPOTHETICAL 45.0 KD P	2.87e+03
584	4	13.8	314	1	Y4AF_RHISN	HYPOTHETICAL 34.7 KD P	2.87e+03	657	4	13.8	387	1	Y4PF_RHISN	PUTATIVE TRANSPOSASE Y	2.87e+03
585	4	13.8	315	1	YK04_CAEEL	HYPOTHETICAL 35.9 KD P	2.87e+03	658	4	13.8	388	1	Y0A5_MYCTU	HYPOTHETICAL 40.2 KD P	2.87e+03
586	4	13.8	315	1	Y775_HAEIN	HYPOTHETICAL TRANSCRIPT	2.87e+03	659	4	13.8	390	1	Y109_NPVAC	HYPOTHETICAL 44.8 KD P	2.87e+03
587	4	13.8	316	1	YU25_HSV6U	U25 PROTEIN	2.87e+03	660	4	13.8	390	1	Y109_NPVOP	HYPOTHETICAL 44.3 KD P	2.87e+03
588	4	13.8	316	1	YVBE_BACSU	HYPOTHETICAL TRANSCRIPT	2.87e+03	661	4	13.8	392	1	Y154_MYCLE	HYPOTHETICAL 42.2 KD P	2.87e+03
589	4	13.8	317	1	Y184_MYCGE	HYPOTHETICAL ADENINE-S	2.87e+03	662	4	13.8	392	1	YAY1_SCHPO	HYPOTHETICAL 44.8 KD P	2.87e+03
590	4	13.8	318	1	YK03_CAEEL	HYPOTHETICAL 36.5 KD P	2.87e+03	663	4	13.8	393	1	Y100_SYNY3	HYPOTHETICAL 42.7 KD P	2.87e+03
591	4	13.8	319	1	Y028_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	664	4	13.8	393	1	Y130_MYCTU	HYPOTHETICAL 41.9 KD P	2.87e+03
592	4	13.8	320	1	YCF5_ECOLI	HYPOTHETICAL 34.6 KD P	2.87e+03	665	4	13.8	393	1	YXLA_ACTMI	XYLOSE ISOMERASE (EC 5	2.87e+03
593	4	13.8	320	1	YWJD_BACSU	HYPOTHETICAL 36.9 KD P	2.87e+03	666	4	13.8	395	1	YDEE_ECOLI	HYPOTHETICAL 42.7 KD P	2.87e+03
594	4	13.8	321	1	YD57_SCHPO	PROBABLE OXIDOREDUCTAS	2.87e+03	667	4	13.8	395	1	YVFI_BPP2	MAJOR TAIL SHEATH PROT	2.87e+03
595	4	13.8	321	1	YH28_YEAST*	HYPOTHETICAL 36.5 KD P	2.87e+03	668	4	13.8	397	1	YEDS_ECOLI	POTENTIAL OUTER MEMBRA	2.87e+03
596	4	13.8	323	1	YKJ2_CAEEL	HYPOTHETICAL 36.9 KD P	2.87e+03	669	4	13.8	399	1	YCFU_ECOLI	HYPOTHETICAL 43.3 KD P	2.87e+03
597	4	13.8	326	1	YKX2_CAEEL	HYPOTHETICAL 34.6 KD P	2.87e+03	670	4	13.8	402	1	YD61_YEAST	HYPOTHETICAL 47.2 KD P	2.87e+03
598	4	13.8	326	1	YAN5_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	671	4	13.8	404	1	YXNM_RHISN	HYPOTHETICAL TRANSPORT	2.87e+03
599	4	13.8	326	1	YAN5_SCHPO	HYPOTHETICAL 37.9 KD P	2.87e+03	672	4	13.8	404	1	YG3X_YEAST	HYPOTHETICAL 46.8 KD P	2.87e+03
600	4	13.8	326	1	VT2_MYXVL	TUMOR NECROSIS FACTOR	2.87e+03	673	4	13.8	406	1	YG43_YEAST	HYPOTHETICAL 47.3 KD P	2.87e+03
601	4	13.8	326	1	Y09G_ROTHX	GLYCOPROTEIN VP7 (SERO	2.87e+03	674	4	13.8	407	1	Y14N_MYCTU	HYPOTHETICAL 44.7 KD P	2.87e+03
602	4	13.8	327	1	YH04_YEAST	PROBABLE OXIDOREDUCTAS	2.87e+03	675	4	13.8	409	1	Y14B_MYCGE	HYPOTHETICAL PROTEIN I	2.87e+03
603	4	13.8	329	1	YAOB_SCHPO	HYPOTHETICAL 37.6 KD P	2.87e+03	676	4	13.8	411	1	YPEB_SYNP2	HYPOTHETICAL PROTEIN M	2.87e+03
604	4	13.8	331	1	Y244_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	677	4	13.8	415	1	Y012_YEAST	HYPOTHETICAL 47.6 KD P	2.87e+03
605	4	13.8	331	1	YJUV3_YEAST	HYPOTHETICAL 35.9 KD P	2.87e+03	678	4	13.8	418	1	YRAB_YEAST	HYPOTHETICAL 49.0 KD P	2.87e+03
606	4	13.8	332	1	YFHB_ECOLI	HYPOTHETICAL 36.4 KD P	2.87e+03	679	4	13.8	418	1	YEB7_SCHPO	HYPOTHETICAL 46.4 KD P	2.87e+03
607	4	13.8	335	1	YHBW_ECOLI	HYPOTHETICAL 37.1 KD P	2.87e+03	680	4	13.8	419	1	YMD2_CAEEL	HYPOTHETICAL 47.3 KD P	2.87e+03

681	4	13.8	420	1	Y103_SYN3	HYPOTHETICAL 45.8 KD P	2.87e+03	754	505	1	YEP4_YEAST	HYPOTHETICAL 56.6 KD P	2.87e+03
682	4	13.8	421	1	Y447_MYCLE	HYPOTHETICAL 45.9 KD P	2.87e+03	755	507	1	Y108_SYN3	PUTATIVE AMMONIUM TRAN	2.87e+03
683	4	13.8	421	1	YG70_YEAST	HYPOTHETICAL 48.3 KD P	2.87e+03	756	509	1	Y96F_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03
684	4	13.8	423	1	YEF8_SCHPO	HYPOTHETICAL 48.0 KD P	2.87e+03	757	510	1	YML4_YEAST	HYPOTHETICAL 58.0 KD P	2.87e+03
685	4	13.8	426	1	WECF_SALTY	PROBABLE 4-ALPHA-L-FUC	2.87e+03	758	510	1	YFK9_YEAST	HYPOTHETICAL 59.4 KD P	2.87e+03
686	4	13.8	429	1	Y179_ARATH	HYPOTHETICAL 47.9 KD P	2.87e+03	759	515	1	YD56_SCHPO	HYPOTHETICAL 59.0 KD P	2.87e+03
687	4	13.8	431	1	Y421_SCHPO	HYPOTHETICAL 49.6 KD P	2.87e+03	760	516	1	YPO2_YEAST	HYPOTHETICAL 57.7 KD P	2.87e+03
688	4	13.8	433	1	YBL4_YEAST	HYPOTHETICAL 50.8 KD P	2.87e+03	761	516	1	YCX4_EUGGR	HYPOTHETICAL 64.3 KD P	2.87e+03
689	4	13.8	438	1	YCC0_YEAST	TRANSPOSON TY1-17 PROT	2.87e+03	762	519	1	Y711_CAEEL	HYPOTHETICAL 59.4 KD P	2.87e+03
690	4	13.8	439	1	XYLA_CLOTS	XYLOSE ISOMERASE (EC 5	2.87e+03	763	519	1	YF6_WTV	STRUCTURAL PROTEIN P6	2.87e+03
691	4	13.8	440	1	YJ26_YEAST	TRANSPOSON TY1 PROTEIN	2.87e+03	764	520	1	YAE_SCHPO	HYPOTHETICAL 58.4 KD P	2.87e+03
692	4	13.8	440	1	Y433_SCHPO	HYPOTHETICAL 49.6 KD P	2.87e+03	765	520	1	WASP_MOUSE	WISKOTT-ALDRICH SYNDRO	2.87e+03
693	4	13.8	441	1	YXEK_BACSU	HYPOTHETICAL 49.3 KD P	2.87e+03	766	520	1	YAH5_SCHPO	VERY HYPOTHETICAL 52.7	2.87e+03
694	4	13.8	441	1	XYLA_BACSP	XYLOSE ISOMERASE (EC 5	2.87e+03	767	520	1	YB9_YEAST	HYPOTHETICAL 59.4 KD P	2.87e+03
695	4	13.8	441	1	XYLA_BACST	XYLOSE ISOMERASE (EC 5	2.87e+03	768	525	1	YBF4_YEAST	HYPOTHETICAL 59.2 KD P	2.87e+03
696	4	13.8	442	1	Y141_ECOLI	INSERTION ELEMENT IS4	2.87e+03	769	528	1	YEP0_YEAST	HYPOTHETICAL 58.1 KD P	2.87e+03
697	4	13.8	442	1	YNB1_YEAST	HYPOTHETICAL 49.9 KD P	2.87e+03	770	529	1	Y243_ARCFU	HYPOTHETICAL PROTEIN A	2.87e+03
698	4	13.8	444	1	YJG0_YEAST	HYPOTHETICAL 50.0 KD P	2.87e+03	771	529	1	YABD_SCHPO	HYPOTHETICAL 58.6 KD P	2.87e+03
699	4	13.8	444	1	YAV5_YANCV	HYPOTHETICAL 48.0 KD P	2.87e+03	772	532	1	YRR5_CAEEL	HYPOTHETICAL 59.8 KD P	2.87e+03
700	4	13.8	449	1	Y753_SYN3	PROBABLE 4-ALPHA-L-FUC	2.87e+03	773	534	1	YX22_CAEEL	HYPOTHETICAL 62.3 KD P	2.87e+03
701	4	13.8	450	1	WECF_ECOLI	HYPOTHETICAL 51.0 KD P	2.87e+03	774	535	1	YRK_CHICK	PROTO-ONCOGENE TYROSIN	2.87e+03
702	4	13.8	451	1	YQ81_CAEEL	HYPOTHETICAL 51.0 KD P	2.87e+03	775	536	1	ZPB_PIG	ZONA PELLUCIDA SPERM-B	2.87e+03
703	4	13.8	451	1	YB27_YEAST	HYPOTHETICAL 50.9 KD P	2.87e+03	776	537	1	YES_XENLA	PROTO-ONCOGENE TYROSIN	2.87e+03
704	4	13.8	453	1	YAW6_SCHPO	HYPOTHETICAL 50.0 KD P	2.87e+03	777	539	1	YES_CANFA	PROTO-ONCOGENE TYROSIN	2.87e+03
705	4	13.8	453	1	Y076_SYN3	HYPOTHETICAL 50.0 KD P	2.87e+03	778	540	1	ZPB_RABIT	ZONA PELLUCIDA SPERM-B	2.87e+03
706	4	13.8	453	1	YOTA_CAEEL	HYPOTHETICAL 51.8 KD P	2.87e+03	779	541	1	YES_MOUSE	PROTO-ONCOGENE TYROSIN	2.87e+03
707	4	13.8	455	1	YMI8_YEAST	HYPOTHETICAL 51.7 KD P	2.87e+03	780	541	1	YES_CHICK	PROTO-ONCOGENE TYROSIN	2.87e+03
708	4	13.8	456	1	YD45_SCHPO	HYPOTHETICAL 50.5 KD P	2.87e+03	781	542	1	ZYX_CHICK	ZYXIN	2.87e+03
709	4	13.8	457	1	YCA7_SYN3	HYPOTHETICAL 52.7 KD P	2.87e+03	782	542	1	Y159_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
710	4	13.8	457	1	YI90_YEAST	HYPOTHETICAL 50.8 KD P	2.87e+03	783	543	1	YES_HUMAN	PROTO-ONCOGENE TYROSIN	2.87e+03
711	4	13.8	458	1	YMT7_YEAST	HYPOTHETICAL 50.5 KD P	2.87e+03	784	543	1	Y274_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
712	4	13.8	458	1	YMS1_YEAST	HYPOTHETICAL 52.2 KD P	2.87e+03	785	544	1	YAGW_ECOLI	HYPOTHETICAL 60.0 KD P	2.87e+03
713	4	13.8	459	1	Y226_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	786	547	1	YAGW_ECOLI	HYPOTHETICAL 60.0 KD P	2.87e+03
714	4	13.8	461	1	Y061_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	787	548	1	Y24C_MYCTU	HYPOTHETICAL 57.9 KD P	2.87e+03
715	4	13.8	461	1	Y896_SYN3	HYPOTHETICAL 49.7 KD P	2.87e+03	788	549	1	X92_TRYBB	PROTEIN X92	2.87e+03
716	4	13.8	462	1	YMA5_YEAST	HYPOTHETICAL 52.7 KD P	2.87e+03	789	551	1	YIDE_HAEIN	HYPOTHETICAL PROTEIN H	2.87e+03
717	4	13.8	463	1	YNAJ_BACSU	HYPOTHETICAL 51.9 KD P	2.87e+03	790	556	1	Y281_MYCGE	HYPOTHETICAL 64.9 KD P	2.87e+03
718	4	13.8	463	1	YGL1_YEAST	HYPOTHETICAL 51.9 KD P	2.87e+03	791	558	1	YQ2B_YEAST	HYPOTHETICAL 59.4 KD P	2.87e+03
719	4	13.8	464	1	YKMS_PSEAE	HYPOTHETICAL PROTEIN I	2.87e+03	792	558	1	YBAL_ECOLI	HYPOTHETICAL 59.4 KD P	2.87e+03
720	4	13.8	467	1	WCAM_SALTY	COLANIC ACID BIOSYNTH	2.87e+03	793	559	1	YFC3_CAEEL	PUTATIVE ABC TRANSPORT	2.87e+03
721	4	13.8	470	1	YSS2_CAEEL	PUTATIVE SERINE CARBOX	2.87e+03	794	560	1	YK5_YEAST	HYPOTHETICAL 63.9 KD P	2.87e+03
722	4	13.8	471	1	Y148_HUMAN	HYPOTHETICAL PROTEIN K	2.87e+03	795	561	1	YEHU_ECOLI	HYPOTHETICAL 62.1 KD P	2.87e+03
723	4	13.8	471	1	Y782_SYN3	HYPOTHETICAL 51.4 KD P	2.87e+03	796	564	1	ZYX_MOUSE	ZYXIN	2.87e+03
724	4	13.8	474	1	Y294_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03	797	569	1	YAR3_SCHPO	HYPOTHETICAL 61.1 KD P	2.87e+03
725	4	13.8	474	1	YTD8_HUMAN	VITAMIN D-BINDING PROT	2.87e+03	798	570	1	YNE3_CAEEL	HYPOTHETICAL 64.2 KD P	2.87e+03
726	4	13.8	474	1	VP61_NPVOP	61 KD PROTEIN HOMOLOG	2.87e+03	799	570	1	ZPB_FELCA	ZONA PELLUCIDA SPERM-B	2.87e+03
727	4	13.8	475	1	Y294_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03	800	572	1	ZYX_HUMAN	ZYXIN (ZYXIN 2)	2.87e+03
728	4	13.8	475	1	YMD8_YEAST	HYPOTHETICAL 52.6 KD P	2.87e+03	801	573	1	YBP4_YEAST	HYPOTHETICAL 64.3 KD P	2.87e+03
729	4	13.8	475	1	YDF6_SCHPO	HYPOTHETICAL 51.8 KD P	2.87e+03	802	575	1	YH26_YEAST	HYPOTHETICAL 65.3 KD P	2.87e+03
730	4	13.8	476	1	YTD8_RABIT	VITAMIN D-BINDING PROT	2.87e+03	803	576	1	Y134_HUMAN	PUTATIVE ATP-DEPENDENT	2.87e+03
731	4	13.8	476	1	YTD8_RAT	VITAMIN D-BINDING PROT	2.87e+03	804	576	1	YQ06_CAEEL	HYPOTHETICAL 52.2 KD P	2.87e+03
732	4	13.8	477	1	YMN2_CAEEL	HYPOTHETICAL ACETYLCHO	2.87e+03	805	576	1	Y152_YEAST	HYPOTHETICAL 64.0 KD P	2.87e+03
733	4	13.8	479	1	Y098_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03	806	577	1	Y171_CAEEL	HYPOTHETICAL 64.9 KD P	2.87e+03
734	4	13.8	480	1	YB90_YEAST	HYPOTHETICAL 55.5 KD P	2.87e+03	807	578	1	YDV8_SCHPO	HYPOTHETICAL 67.9 KD P	2.87e+03
735	4	13.8	481	1	Y588_ARCFU	HYPOTHETICAL PROTEIN A	2.87e+03	808	581	1	YD15_SCHPO	PUTATIVE REGULATORY PR	2.87e+03
736	4	13.8	481	1	YAB8_SCHPO	PUTATIVE ATP-DEPENDENT	2.87e+03	809	585	1	YKJ0_YEAST	HYPOTHETICAL 67.5 KD P	2.87e+03
737	4	13.8	481	1	XYLB_STRRU	XYLOSE KINASE (EC 2	2.87e+03	810	587	1	YK00_YEAST	HYPOTHETICAL 67.5 KD P	2.87e+03
738	4	13.8	482	1	YVSC_TRYBB	VARIANT SURFACE GLYCOP	2.87e+03	811	588	1	YK06_CAEEL	HYPOTHETICAL 67.3 KD P	2.87e+03
739	4	13.8	484	1	YAV5_SCHPO	PROBABLE AMINOTRANSFER	2.87e+03	812	592	1	YK06_CAEEL	HYPOTHETICAL 66.9 KD P	2.87e+03
740	4	13.8	488	1	YDP8_SCHPO	HYPOTHETICAL 54.2 KD T	2.87e+03	813	598	1	YAUU_SCHPO	HYPOTHETICAL 60.7 KD P	2.87e+03
741	4	13.8	489	1	YABN_BACSU	HYPOTHETICAL 56.1 KD P	2.87e+03	814	599	1	YJUI_YEAST	HYPOTHETICAL 68.5 KD P	2.87e+03
742	4	13.8	490	1	YIHF_ECOLI	HYPOTHETICAL 54.2 KD P	2.87e+03	815	600	1	YB57_SCHPO	HYPOTHETICAL 69.4 KD P	2.87e+03
743	4	13.8	491	1	Y225_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03	816	602	1	YH02_YEAST	HYPOTHETICAL 69.0 KD P	2.87e+03
744	4	13.8	491	1	YCT5_YEAST	HYPOTHETICAL SUBTILASE	2.87e+03	817	605	1	WIS1_SCHPO	PROTEIN KINASE WIS1 (E	2.87e+03
745	4	13.8	493	1	XYLB_HAEIN	XYLOSE KINASE (EC 2	2.87e+03	818	606	1	Y4B0_RHNSN	HYPOTHETICAL 66.8 KD P	2.87e+03
746	4	13.8	493	1	VPE_VTCSA	VACUOLAR PROCESSING EN	2.87e+03	819	611	1	YD3M_HERAU	HYPOTHETICAL 68.4 KD P	2.87e+03
747	4	13.8	497	1	YRICO	VACUOLAR PROCESSING EN	2.87e+03	820	611	1	YCR3_YEAST	HYPOTHETICAL 69.2 KD P	2.87e+03
748	4	13.8	502	1	Y190_HELPY	HYPOTHETICAL PROTEIN H	2.87e+03	821	611	1	YJS1_YEAST	HYPOTHETICAL 69.5 KD P	2.87e+03
749	4	13.8	502	1	WASKOTT-ALDRICH SYNDRO	2.87e+03	822	611	1	YX45_MYCTU	HYPOTHETICAL 67.2 KD P	2.87e+03	
750	4	13.8	503	1	YP88_CAEEL	PUTATIVE GLUCOSYLTRANS	2.87e+03	823	612	1	YNB8_YEAST	HYPOTHETICAL 69.6 KD P	2.87e+03
751	4	13.8	503	1	YN23_YEAST	HYPOTHETICAL 52.3 KD P	2.87e+03	824	614	1	YDNK_LACLA	HYPOTHETICAL 70.0 KD P	2.87e+03
752	4	13.8	503	1	YSA1_TRYBB	VARIANT SURFACE GLYCOP	2.87e+03	825	614	1	ZF29_MOUSE	ZINC FINGER PROTEIN 29	2.87e+03
753	4	13.8	504	1	YM68_YEAST	HYPOTHETICAL 56.2 KD P	2.87e+03	826	614	1	YQ3F_YEAST	HYPOTHETICAL 67.6 KD P	2.87e+03

827	4	13.8	615	1	YK29_YEAST	HYPOTHETICAL 68.9 KD P	2.87e+03	900	4	13.8	888	1	YJH0_YEAST	HYPOTHETICAL 104.3 KD	2.87e+03
828	4	13.8	616	1	YJGL_ECOLI	HYPOTHETICAL 71.1 KD P	2.87e+03	901	4	13.8	888	1	YGB4_YEAST	HYPOTHETICAL REGULATOR	2.87e+03
829	4	13.8	617	1	Y4A2_HAEN	HYPOTHETICAL PROTEIN H	2.87e+03	902	4	13.8	908	1	VP2_BRD	STRUCTURAL CORE PROTEI	2.87e+03
830	4	13.8	620	1	YM20_YEAST	HYPOTHETICAL 72.2 KD P	2.87e+03	903	4	13.8	917	1	YJ33_YEAST	HYPOTHETICAL 104.8 KD	2.87e+03
831	4	13.8	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	2.87e+03	904	4	13.8	918	1	YNE6_CAEEL	PROBABLE 3',5'-CYCLIC	2.87e+03
832	4	13.8	620	1	YAK8_SCHPO	PUTATIVE MULTICOPPER O	2.87e+03	905	4	13.8	919	1	YK05_YEAST	HYPOTHETICAL 105.7 KD	2.87e+03
833	4	13.8	630	1	Y4B0_RHISN	HYPOTHETICAL 67.9 KD P	2.87e+03	906	4	13.8	920	1	YML1_YEAST	HYPOTHETICAL 104.8 KD	2.87e+03
834	4	13.8	640	1	Y551_SVNY3	HYPOTHETICAL 70.4 KD P	2.87e+03	907	4	13.8	921	1	YEG9_SCHPO	HYPOTHETICAL 104.4 KD	2.87e+03
835	4	13.8	641	1	YGG0_YEAST	HYPOTHETICAL 73.1 KD P	2.87e+03	908	4	13.8	923	1	YEL4_YEAST	HYPOTHETICAL 103.3 KD	2.87e+03
836	4	13.8	645	1	Y002_MYCTU	HYPOTHETICAL 69.2 KD P	2.87e+03	909	4	13.8	923	1	YAU4_SCHPO	HYPOTHETICAL 104.5 KD	2.87e+03
837	4	13.8	645	1	Y081_CAEEL	HYPOTHETICAL GSP-BINDI	2.87e+03	910	4	13.8	925	1	YNOA_CAEEL	HYPOTHETICAL 105.5 KD	2.87e+03
838	4	13.8	646	1	YD49_SCHPO	HYPOTHETICAL 76.2 KD P	2.87e+03	911	4	13.8	928	1	Y233_YEAST	HYPOTHETICAL 106.7 KD	2.87e+03
839	4	13.8	646	1	WEEL_HUMAN	WEEL-LIKE PROTEIN KINA	2.87e+03	912	4	13.8	932	1	YAI4_SCHPO	HYPOTHETICAL 103.2 KD	2.87e+03
840	4	13.8	646	1	WEEL_MOUSE	WEEL-LIKE PROTEIN KINA	2.87e+03	913	4	13.8	943	1	YMI1_YEAST	HYPOTHETICAL 105.9 KD	2.87e+03
841	4	13.8	654	1	YED1_SCHPO	HYPOTHETICAL 73.9 KD P	2.87e+03	914	4	13.8	944	1	YQ35_YEAST	VACUOLAR PROTEIN SORTI	2.87e+03
842	4	13.8	655	1	YJ68_YEAST	HYPOTHETICAL 74.1 KD P	2.87e+03	915	4	13.8	953	1	YCP7_YEAST	OUTER CAPSID PROTEIN V	2.87e+03
843	4	13.8	661	1	YN84_YEAST	HYPOTHETICAL 75.4 KD P	2.87e+03	916	4	13.8	956	1	VP2_BTIV10	OUTER CAPSID PROTEIN V	2.87e+03
844	4	13.8	661	1	YJCS_ECOLI	HYPOTHETICAL 73.2 KD P	2.87e+03	917	4	13.8	959	1	VP2_BTIV3V	OUTER CAPSID PROTEIN V	2.87e+03
845	4	13.8	662	1	YFL8_YEAST	HYPOTHETICAL 75.9 KD P	2.87e+03	918	4	13.8	962	1	YAXA_SCHPO	HYPOTHETICAL 112.7 KD	2.87e+03
846	4	13.8	662	1	YME1_SCHMA	YME1 PROTEIN HOMOLOG (	2.87e+03	919	4	13.8	962	1	Y4AE_SCHPO	HYPOTHETICAL 107.3 KD	2.87e+03
847	4	13.8	664	1	YMC0_YEAST	HYPOTHETICAL 76.1 KD P	2.87e+03	920	4	13.8	964	1	YINO_YEAST	PUTATIVE 108.8 KD TRAN	2.87e+03
848	4	13.8	669	1	YPH_BPP2	PROBABLE TAIL FIBRE PR	2.87e+03	921	4	13.8	971	1	YQIA_CAEEL	HYPOTHETICAL 111.8 KD	2.87e+03
849	4	13.8	673	1	YNM5_YEAST	HYPOTHETICAL 73.8 KD P	2.87e+03	922	4	13.8	993	1	YD43_SCHPO	HYPOTHETICAL 111.9 KD	2.87e+03
850	4	13.8	676	1	Y043_YEAST	HYPOTHETICAL 77.0 KD P	2.87e+03	923	4	13.8	995	1	YIO9_YEAST	HYPOTHETICAL 99.7 KD P	2.87e+03
851	4	13.8	679	1	YI83_YEAST	HYPOTHETICAL 74.8 KD P	2.87e+03	924	4	13.8	996	1	YHN3_YEAST	HYPOTHETICAL 113.8 KD	2.87e+03
852	4	13.8	682	1	YEC3_YEAST	HYPOTHETICAL 78.3 KD P	2.87e+03	925	4	13.8	997	1	YNM3_YEAST	HYPOTHETICAL 110.9 KD	2.87e+03
853	4	13.8	682	1	YI24_HUMAN	HYPOTHETICAL PROTEIN K	2.87e+03	926	4	13.8	1000	1	Y083_CAEEL	HYPOTHETICAL 113.4 KD	2.87e+03
854	4	13.8	683	1	YJQ0_ECOLI	HYPOTHETICAL 76.1 KD P	2.87e+03	927	4	13.8	1004	1	YG21_YEAST	HYPOTHETICAL 113.9 KD	2.87e+03
855	4	13.8	687	1	YL18_CAEEL	HYPOTHETICAL 76.7 KD P	2.87e+03	928	4	13.8	1005	1	Y321_MYCPN	HYPOTHETICAL LIPOPROTE	2.87e+03
856	4	13.8	689	1	YVAL_BACSU	HYPOTHETICAL 78.8 KD P	2.87e+03	929	4	13.8	1008	1	YQ2N_CAEEL	PUTATIVE PRE-MRNA SPLI	2.87e+03
857	4	13.8	692	1	YJH3_YEAST	HYPOTHETICAL 80.4 KD P	2.87e+03	930	4	13.8	1010	1	WNT5_DROME	PROTEIN DWNT-5 PRECURS	2.87e+03
858	4	13.8	694	1	Y4BA_RHISN	HYPOTHETICAL 78.7 KD P	2.87e+03	931	4	13.8	1010	1	YKK1_CAEEL	HYPOTHETICAL 113.2 KD	2.87e+03
859	4	13.8	701	1	Y185_MYCGE	HYPOTHETICAL LIPOPROTE	2.87e+03	932	4	13.8	1030	1	Y075_MYCPN	HYPOTHETICAL PROTEIN M	2.87e+03
860	4	13.8	703	1	Y4B1_RHISN	HYPOTHETICAL 76.2 KD P	2.87e+03	933	4	13.8	1030	1	VPPI_CAEEL	PUTATIVE CLATHRIN-COAT	2.87e+03
861	4	13.8	711	1	YCBY_HAEN	HYPOTHETICAL PROTEIN H	2.87e+03	934	4	13.8	1032	1	Y895_HAEN	HYPOTHETICAL CLATHRIN-COAT	2.87e+03
862	4	13.8	714	1	YJL2_YEAST	HYPOTHETICAL 80.0 KD T	2.87e+03	935	4	13.8	1037	1	YQJ8_YEAST	HYPOTHETICAL 118.4 KD	2.87e+03
863	4	13.8	720	1	YCG5_ECOLI	HYPOTHETICAL 82.0 KD P	2.87e+03	936	4	13.8	1039	1	Y4JQ_RHISN	HYPOTHETICAL 115.9 KD	2.87e+03
864	4	13.8	722	1	YG42_YEAST	HYPOTHETICAL 78.8 KD P	2.87e+03	937	4	13.8	1044	1	YDE1_SCHPO	HYPOTHETICAL PROTEIN C	2.87e+03
865	4	13.8	733	1	YFJ2_YEAST	HYPOTHETICAL 79.7 KD P	2.87e+03	938	4	13.8	1060	1	YN18_YEAST	HYPOTHETICAL 118.3 KD	2.87e+03
866	4	13.8	751	1	VP4_ROTGI	OUTER CAPSID PROTEIN V	2.87e+03	939	4	13.8	1085	1	YAP4_SCHPO	HYPOTHETICAL 122.9 KD	2.87e+03
867	4	13.8	751	1	Y226_HUMAN	HYPOTHETICAL PROTEIN K	2.87e+03	940	4	13.8	1093	1	YKD5_CAEEL	PUTATIVE SERINE/THREON	2.87e+03
868	4	13.8	753	1	YJ03_CAEEL	HYPOTHETICAL 85.1 KD N	2.87e+03	941	4	13.8	1107	1	YAX5_SCHPO	HYPOTHETICAL 126.5 KD	2.87e+03
869	4	13.8	756	1	YDLF_SCHPO	HYPOTHETICAL 87.5 KD P	2.87e+03	942	4	13.8	1110	1	YNO3_YEAST	HYPOTHETICAL 124.5 KD	2.87e+03
870	4	13.8	764	1	YJQ0_YEAST	HYPOTHETICAL 87.2 KD P	2.87e+03	943	4	13.8	1118	1	YIP1_YEAST	HYPOTHETICAL 128.7 KD	2.87e+03
871	4	13.8	775	1	YHGE_BACSU	HYPOTHETICAL 84.1 KD P	2.87e+03	944	4	13.8	1130	1	YQ38_CAEEL	HYPOTHETICAL 130.5 KD	2.87e+03
872	4	13.8	778	1	YV05_METJA	PUTATIVE ATP-DEPENDENT	2.87e+03	945	4	13.8	1134	1	YML7_YEAST	HYPOTHETICAL 126.1 KD	2.87e+03
873	4	13.8	782	1	YV02_YEAST	HYPOTHETICAL 88.8 KD P	2.87e+03	946	4	13.8	1139	1	Y050_CAEEL	HYPOTHETICAL 128.6 KD	2.87e+03
874	4	13.8	783	1	YAYE_SCHPO	HYPOTHETICAL 89.2 KD P	2.87e+03	947	4	13.8	1159	1	YAJ4_SCHPO	HYPOTHETICAL 123.9 KD	2.87e+03
875	4	13.8	787	1	YKEL_DROME	HYPOTHETICAL 85.6 KD P	2.87e+03	948	4	13.8	1165	1	YNF4_YEAST	HYPOTHETICAL 128.1 KD	2.87e+03
876	4	13.8	792	1	Y85K_SSV1	HYPOTHETICAL 85.7 KD P	2.87e+03	949	4	13.8	1170	1	YKD8_YEAST	PUTATIVE 128.2 KD TRAN	2.87e+03
877	4	13.8	796	1	YAB9_SCHPO	HYPOTHETICAL 90.9 KD P	2.87e+03	950	4	13.8	1188	1	YEC5_YEAST	HYPOTHETICAL 136.1 KD	2.87e+03
878	4	13.8	796	1	YFC3_YEAST	HYPOTHETICAL 91.4 KD P	2.87e+03	951	4	13.8	1191	1	Y143_SVNY3	HYPOTHETICAL WD-REPEAT	2.87e+03
879	4	13.8	801	1	YV34_SCHPO	PHOSPHATIDYLINOSITOL 3	2.87e+03	952	4	13.8	1203	1	XCPE_XENLA	CHROMOSOME ASSEMBLY PR	2.87e+03
880	4	13.8	801	1	YEV7_YEAST	HYPOTHETICAL 92.5 KD P	2.87e+03	953	4	13.8	1204	1	YAEF_SCHPO	HYPOTHETICAL LIPOPROTE	2.87e+03
881	4	13.8	802	1	YGN9_YEAST	HYPOTHETICAL 90.8 KD P	2.87e+03	954	4	13.8	1225	1	Y309_MYCGE	HYPOTHETICAL LIPOPROTE	2.87e+03
882	4	13.8	804	1	VP5_MTV	OUTER COAT PROTEIN P5	2.87e+03	955	4	13.8	1240	1	YNI1_YEAST	HYPOTHETICAL 141.5 KD	2.87e+03
883	4	13.8	819	1	Y083_NVP0P	HYPOTHETICAL 91.1 KD P	2.87e+03	956	4	13.8	1244	1	Y307_MYCPN	HYPOTHETICAL LIPOPROTE	2.87e+03
884	4	13.8	831	1	VPPI_HUMAN	CLATHRIN-COATED VESICL	2.87e+03	957	4	13.8	1246	1	YPO3_HSVSA	PROBABLE MEMBRANE ANTI	2.87e+03
885	4	13.8	834	1	YPM2_YEAST	CLATHRIN-COATED VESICL	2.87e+03	958	4	13.8	1253	1	YAGD_SCHPO	HYPOTHETICAL 143.3 KD	2.87e+03
886	4	13.8	838	1	VPPI_BOVIN	CLATHRIN-COATED VESICL	2.87e+03	959	4	13.8	1254	1	YNCA_CAEEL	HYPOTHETICAL 142.5 KD	2.87e+03
887	4	13.8	839	1	YME1_YEAST	HYPOTHETICAL 97.0 KD P	2.87e+03	960	4	13.8	1257	1	YK28_CAEEL	HYPOTHETICAL 143.2 KD	2.87e+03
888	4	13.8	843	1	YMS1_YEAST	HYPOTHETICAL 93.3 KD P	2.87e+03	961	4	13.8	1266	1	Y468_MYCGE	HYPOTHETICAL PROTEIN M	2.87e+03
889	4	13.8	844	1	YD25_SCHPO	HYPOTHETICAL 94.9 KD P	2.87e+03	962	4	13.8	1272	1	Y228_METJA	HYPOTHETICAL PROTEIN M	2.87e+03
890	4	13.8	847	1	Y083_NPVAC	HYPOTHETICAL 96.2 KD P	2.87e+03	963	4	13.8	1273	1	WEB1_YEAST	WEB1 PROTEIN (PROTEIN	2.87e+03
891	4	13.8	855	1	Y050_MYCTU	PUTATIVE CATION-TRANSP	2.87e+03	964	4	13.8	1273	1	YAR2_SCHPO	HYPOTHETICAL 142.5 KD	2.87e+03
892	4	13.8	856	1	YPH1_NEUCR	VACUOLAR ATP SYNTHASE	2.87e+03	965	4	13.8	1309	1	YAB2_SCHPO	HYPOTHETICAL 150.5 KD	2.87e+03
893	4	13.8	863	1	YEJO_ECOLI	HYPOTHETICAL 91.2 KD P	2.87e+03	966	4	13.8	1325	1	Y309_MYCPN	HYPOTHETICAL LIPOPROTE	2.87e+03
894	4	13.8	878	1	YD97_SCHPO	HYPOTHETICAL 98.4 KD P	2.87e+03	967	4	13.8	1325	1	YAE6_SCHPO	HYPOTHETICAL 145.8 KD	2.87e+03
895	4	13.8	881	1	YJTB_YEAST	HYPOTHETICAL 97.7 KD M	2.87e+03	968	4	13.8	1328	1	YME4_YEAST	TRANSPOSON TY1 PROTEIN	2.87e+03
896	4	13.8	883	1	YB3D_SCHPO	PUTATIVE KINESIN-LIKE	2.87e+03	969	4	13.8	1328	1	YMT5_YEAST	TRANSPOSON TY1 PROTEIN	2.87e+03
897	4	13.8	886	1	YMR1_CAEEL	PUTATIVE ATP-DEPENDENT	2.87e+03	970	4	13.8	1328	1	YMD9_YEAST	PHAGE-LIKE ELEMENT PBS	2.87e+03
898	4	13.8	886	1	YEB6_SCHPO	HYPOTHETICAL 98.3 KD P	2.87e+03	971	4	13.8	1332	1	YKDO_BACSU	PROBABLE ATP-DEPENDENT	2.87e+03
899	4	13.8	887	1	YAV5_SCHPO	PROBABLE ATP-DEPENDENT	2.87e+03	972	4	13.8	1333	1	YN99_YEAST	PROBABLE ATP-DEPENDENT	2.87e+03

```

973 4 13.8 1351 1 YOL4_CAEEL HYPOTHETICAL 152.4 KD 2.87e+03
974 4 13.8 1363 1 XDH_EMENI XANTHINE DEHYDROGENASE 2.87e+03
975 4 13.8 1403 1 YDF3_SCHPO PROBABLE EUKARYOTIC IN 2.87e+03
976 4 13.8 1420 1 YMBB_YEAST HYPOTHETICAL 163.6 KD 2.87e+03
977 4 13.8 1432 1 WRN_HUMAN WERNER SYNDROME HELICA 2.87e+03
978 4 13.8 1442 1 YJ9F_YEAST HYPOTHETICAL 161.2 KD 2.87e+03
979 4 13.8 1453 1 VP15_YEAST PROTEIN KINASE VPS15 ( 2.87e+03
980 4 13.8 1478 1 YAMB_SCHPO PROBABLE ATP-DEPENDENT 2.87e+03
981 4 13.8 1507 1 Y056_HUMAN HYPOTHETICAL PROTEIN K 2.87e+03
982 4 13.8 1541 1 YHV6_YEAST HYPOTHETICAL 175.8 KD 2.87e+03
983 4 13.8 1584 1 YJ9G_YEAST HYPOTHETICAL 182.0 KD 2.87e+03
984 4 13.8 1636 1 YN37_YEAST HYPOTHETICAL 186.8 KD 2.87e+03
985 4 13.8 1738 1 YCF1_EPIVI HYPOTHETICAL 208 KD PR 2.87e+03
986 4 13.8 1745 1 YL51_CAEEL HYPOTHETICAL PROTEIN F 2.87e+03
987 4 13.8 1755 1 YJ29_YEAST TRANSPOSON TY1 PROTEIN 2.87e+03
988 4 13.8 1756 1 YCF1_PINTH HYPOTHETICAL 205.3 KD 2.87e+03
989 4 13.8 1788 1 YP72_CAEEL HYPOTHETICAL 200.6 KD 2.87e+03
990 4 13.8 1794 1 YAV1_SCHPO HYPOTHETICAL 193.3 KD 2.87e+03
991 4 13.8 1803 1 YJL3_YEAST TRANSPOSON TY4 207.7 K 2.87e+03
992 4 13.8 1868 1 YHD0_YEAST HYPOTHETICAL 210.4 KD 2.87e+03
993 4 13.8 1941 1 YRM8_CAEEL HYPOTHETICAL 216.3 KD 2.87e+03
994 4 13.8 2334 1 WAPA_BACSU WALL-ASSOCIATED PROTEI 2.87e+03
995 4 13.8 2476 1 ZAN_PIG ZONADHESIN PRECURSOR. 2.87e+03
996 4 13.8 2688 1 ZEP1_MOUSE ZINC FINGER PROTEIN 40 2.87e+03
997 4 13.8 3005 1 YFH2_DROME ZINC-FINGER PROTEIN 2 2.87e+03
998 4 13.8 3343 1 YOG7_CAEEL HYPOTHETICAL 375.7 KD 2.87e+03
999 4 13.8 3655 1 YAMB_SCHPO HYPOTHETICAL 420.8 KD 2.87e+03
1000 4 13.8 4385 1 YP73_CAEEL HYPOTHETICAL 486.0 KD 2.87e+03

```

## ALIGNMENTS

```

RESULT 1
ID PLB1 PENNO STANDARD; PRT; 612 AA.
AC P39457;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYSOPHOSPHOLIPASE PRECURSOR (EC 3.1.1.5) (PHOSPHOLIPASE B) (FRAGMENT).
OS PENICILLIUM NOTATUM
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EURASCOMYCETES; PLECTOMYCETES;
OC EUROTIALES; TRICHOCOMACEAE; MITOSPORIC TRICHOCOMACEAE; PENICILLIUM.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 10-20 AND 185-199.
RC STRAIN-INFO 4640;
RX MEDLINE; 92111525;
RA MASUDA N., KITAMURA N., SAITO K.;
RT "Primary structure of protein moiety of Penicillium notatum
phospholipase B deduced from the cDNA.";
RL EUR. J. BIOCHEM. 202:783-787(1991).
CC -!- FUNCTION: CATALYZES THE RELEASE OF FATTY ACIDS FROM
LYSOPHOSPHOLIPIDS.
CC -!- CATALYTIC ACTIVITY: 2-LYSOPHOSPHATIDYLCHOLINE + H(2)O =
GLYCEROPHOSPHOCHOLINE + A FATTY ACID ANION.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- PTM: N-GLYCOSYLATED.
CC -!- SIMILARITY: TO OTHER FUNGAL LYSOPHOSPHOLIPASES AND TO YEAST
SPOL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X60348; G3184;
DR PIR; S29318; S29318.
KW LIPID DEGRADATION; HYDROLASE; GLYCOPROTEIN; SIGNAL.
FT NON_TER 1 1
FT SIGNAL <1 9
FT CHAIN 10 612 LYSOPHOSPHOLIPASE.

```

```

FT DISULFID 16 ?
FT CARBOHYD 41 41
FT CARBOHYD 81 81
FT CARBOHYD 116 116
FT CARBOHYD 150 150
FT CARBOHYD 223 223
FT CARBOHYD 267 267
FT CARBOHYD 306 306
FT CARBOHYD 335 335
FT CARBOHYD 427 427
FT CARBOHYD 440 440
FT CARBOHYD 446 446
FT CARBOHYD 477 477
FT CARBOHYD 498 498
FT CARBOHYD 526 526
FT CARBOHYD 532 532
FT CARBOHYD 567 567
FT CARBOHYD 571 571
SQ SEQUENCE 612 AA; 65750 MW; 16BESBEC CRC32;

Query Match 24.1%; Score 7; DB 1; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.11e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 PSPLVYV 487
| | | | |
QY 15 PSPLVYV 21

RESULT 2
ID BLAC_KLEPN STANDARD; PRT; 279 AA.
AC P05192;
DT 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE BETA-LACTAMASE PRECURSOR (EC 3.5.2.6) (PENICILLINASE).
OS KLEBSIELLA PNEUMONIAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC KLEBSIELLA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LEN-1;
RX MEDLINE; 87030906.
RA ARAKAWA Y., OHTA M., KIDO N., FUJII Y., KOMATSU T., KATO N.;
RT "Close evolutionary relationship between the chromosomally encoded
beta-lactamase gene of Klebsiella pneumoniae and the TEM
beta-lactamase gene mediated by R plasmids.";
RL FEBS LETT. 207:69-74(1986).
CC -!- CATALYTIC ACTIVITY: A BETA-LACTAM + H(2)O = A SUBSTITUTED BETA-
AMINO ACID.
CC -!- SIMILARITY: BELONGS TO THE CLASS-A BETA-LACTAMASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X04515; G43813;
DR PIR; A24469; A24469.
DR PROSITE; PS00146; BETA-LACTAMASE_A; 1.
DR PFAM; PF00144; beta-lactamase; 1.
DR HSP; P00810; IXPB.
KW HYDROLASE; ANTIBIOTIC RESISTANCE; SIGNAL.
FT SIGNAL 1 21
FT CHAIN 22 279 BETA-LACTAMASE.
FT ACT_SITE 66 66 BY SIMILARITY.
FT DISULFID 73 119 BY SIMILARITY.
FT BINDING 230 232 SUBSTRATE (BY SIMILARITY).
SQ SEQUENCE 279 AA; 30269 MW; 957B1FFE CRC32;

```

Query Match 20.7%; Score 6; DB 1; Length 279;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 PLVVYA 21  
| | | | |  
QY 17 PLVVYA 22

RESULT 3  
ID QOX2\_BACSU STANDARD; PRT; 323 AA.  
AC P34957;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE QUINOL OXIDASE POLYPEPTIDE II PRECURSOR (EC 1.9.3.-) (QUINOL OXIDASE  
DE AA3-600, SUBUNIT QOXA) (OXIDASE AA(3) SUBUNIT 2).  
GN QOXA OR IPA-37D.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE: 92268053.  
RA SANTANA M., KUNST F., HULLO M.F., RAPOPORT G., DANCHIN A.,  
RA GLASER P.;  
RT "Molecular cloning, sequencing, and physiological characterization of  
RT the qox operon from *Bacillus subtilis* encoding the aa3-600 quinol  
RT oxidase.";  
RL J. BIOL. CHEM. 267:10225-10231(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE: 95020537.  
RA GLASER P., KUNST F., ARNAUD M., COUDART M.P., GONZALES W.,  
RA HULLO M.F., IONESCU M., LUBOCHINSKY B., MARCELINO L., MOSZER I.,  
RA PRESECAN E., SANTANA M., SCHNEIDER E., SCHWEIZER J., VERTES A.,  
RA RAPOPORT G., DANCHIN A.;  
RT "Bacillus subtilis genome project: cloning and sequencing of the 97  
RT kb region from 325 degrees to 333 degrees.";  
RL MOL. MICROBIOL. 10:371-384(1993)  
CC -!- FUNCTION: CATALYZES QUINOL OXIDATION WITH THE CONCOMITANT  
CC REDUCTION OF OXYGEN TO WATER. MAJOR COMPONENT FOR ENERGY  
CC CONVERSION DURING VEGETATIVE GROWTH. SUBUNIT II TRANSFERS THE  
CC ELECTRONS FROM A QUINOL TO THE BINUCLEAR CENTER OF THE CATALYTIC  
CC SUBUNIT I.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- SIMILARITY: SOME, TO MITOCHONDRIAL OR BACTERIAL COX2 SUBUNITS.  
CC BUT LACK HEME-BINDING DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M86548; G143396; ALT\_INIT.  
CC EMBL; X73124; G580870; ALT\_INIT.  
CC PIR; A38129; A38129.  
CC SUBTILIST; BG10583; QOXA.  
CC PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
CC HSSP; P18400; 1CWV.  
KW OXIDOREDUCTASE; ELECTRON TRANSPORT; TRANSMEMBRANE; SIGNAL;  
LIPOPROTEIN.  
FT SIGNAL 1 26 POTENTIAL.  
FT CHAIN 27 323 QUINOL OXIDASE POLYPEPTIDE II.  
FT LIPID 27 27 N-ACYL DIGLYCERIDE (POTENTIAL).  
FT TRANSMEM 51 71 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
SQ SEQUENCE 323 AA; 36468 MW; AD299C18 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 323;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 132 PLVVYA 137  
| | | | |  
QY 17 PLVVYA 22

RESULT 4  
ID FLAA\_YEAST STANDARD; PRT; 361 AA.  
AC P10664;  
DT 01-JUL-1989 (REL. 11, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE 60S RIBOSOMAL PROTEIN L4-A (L2A) (RP2).  
GN RPL4A OR RPL2A OR RPL2 OR YBR031W OR YBR0315.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 88198160.  
RA PRESUTTI C., LUCIOLI A., BOZZONI I.;  
RT "Ribosomal protein L2 in *Saccharomyces cerevisiae* is homologous to  
RT ribosomal protein L1 in *Xenopus laevis*. Isolation and  
RT characterization of the genes.";  
RL J. BIOL. CHEM. 263:6188-6192(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE: 94378725.  
RA SMITS P.H.M., DE HAAN M., MAAT C., GRIVELL L.A.;  
RT "The complete sequence of a 33 kb fragment on the right arm of  
RT chromosome II from *Saccharomyces cerevisiae* reveals 16 open reading  
RT frames, including ten new open reading frames, five previously  
RT identified genes and a homologue of the SCO1 gene.";  
RN YEAST 10:S75-S80(1994).  
RN [3]  
RP L2 MRNA REGULATION.  
RX MEDLINE: 91293097.  
RA PRESUTTI C., CIAPRE S.-A., BOZZONI I.;  
RT "The ribosomal protein L2 in *S. cerevisiae* controls the level of  
RT accumulation of its own mRNA.";  
RL EMBO J. 10:2215-2221(1991).  
RN [4]  
RP SEQUENCE OF 1-20.  
RX MEDLINE: 92184799.  
RA TAKAKURA H., TSUNASAWA S., MIYAGI M., WARNER J.R.;  
RT "NH2-terminal acetylation of ribosomal proteins of *Saccharomyces*  
RT *cerevisiae*.";  
RL J. BIOL. CHEM. 267:5442-5445(1992).  
CC -!- FUNCTION: PARTICIPATES IN THE REGULATION OF THE ACCUMULATION  
CC OF ITS OWN MRNA.  
CC -!- THERE ARE TWO GENES FOR L4 IN YEAST.  
CC -!- SIMILARITY: BELONGS TO THE L4E FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X76078; G498759; -.  
CC EMBL; J03195; G172413; -.  
CC EMBL; Z35900; G536244; -.  
CC PIR; S45887; S45887.  
CC SGD; L0001701; RPL4A.  
CC PROSITE; PS00939; RIBOSOMAL\_LIE; 1.  
CC PFAM; PF00573; Lie; 1.



2



DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE UBIQUINOL-CYTOCHROME-C REDUCTASE COMPLEX CORE PROTEIN I PRECURSOR  
DE (EC 1.10.2.2).  
GN QCR1 OR COR1 OR YBL045C OR YBL0403.  
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;  
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 87057439.  
RA TZAGOLOFF A., WU M., CRIVELLONE M.;  
RT "Assembly of the mitochondrial membrane system. Characterization of  
RT COR1, the structural gene for the 44-kilodalton core protein of yeast  
RT coenzyme QH2-cytochrome c reductase.";  
RL J. BIOL. CHEM. 261:17163-17169(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RX DE WERFOSSE P., JACQUES B., JONNIAUX J.-L., PURNELLE B., SKALA J.,  
RA GOFFEAU A.;  
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast  
RT chromosome II reveals homologues to bacterial proline synthetase and  
RT murine alpha-adaptin, as well as a new permease and a DNA-binding  
RT protein.";  
RL YEAST 10:1489-1496(1994).  
CC -!- FUNCTION: THIS IS A COMPONENT OF THE UBIQUINOL-CYTOCHROME C  
CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH  
CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY  
CC MEDIATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
CC -!- CATALYTIC ACTIVITY: OH(2) + 2 FERRICYTOCHROME C = O +  
CC 2 FERROCYTOCHROME C.  
CC -!- SUBUNIT: FUNGI BCL COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY  
CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE  
CC INSULINASE FAMILY. DOES NOT SEEM TO HAVE A PROTEASE ACTIVITY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J02636; G171256; -  
DR EMBL: X78214; G463264; -  
DR EMBL: Z35806; G536065; -  
DR PIR: A25351; A25351.  
DR PIR: S42500; S42500.  
DR SGD: L0000381; COR1.  
DR PROSITE: PS00143; INSULINASE; FALSE\_NEG.  
DR PFAM: PF00675; Insulinase; 1.  
DR MITOCHONDRION; INNER MEMBRANE; ELECTRON TRANSPORT; RESPIRATORY CHAIN;  
KW OXIDOREDUCTASE; TRANSIT PEPTIDE.  
FT TRANSIT 1 26  
FT CHAIN 27 457 UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX  
FT CORE PROTEIN 1.  
SQ SEQUENCE 457 AA; 50227 MW; D80FF91C CRC32;  
Query Match 20.7%; Score 6; DB 1; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 4 TVTSKT 9  
QY 2 TVTSKT 7  
RESULT 8  
ID PURA\_CABEL STANDARD: PRT; 457 AA.

AC P91134;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PROBABLE ADENYLOSUCCLINATE SYNTHETASE (EC 6.3.4.4) (IMP--ASPARTATE  
DE LIGASE).  
GN C37H5.6.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA DAVIDSON S., GILLAM B.;  
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE  
CC NUCLEOTIDE BIOSYNTHESIS.  
CC -!- CATALYTIC ACTIVITY: GTP + IMP + L-ASPARTATE = GDP +  
CC ORTHOPHOSPHATE + ADENYLOSUCCLINATE.  
CC -!- PATHWAY: FIRST COMMITTED STEP IN AMP BIOSYNTHESIS.  
CC -!- SIMILARITY: WITH OTHER ADENYLOSUCCLINATE SYNTHETASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U88315; G1825781; -  
DR WORMPEP: C37H5.6; CE08630.  
DR PROSITE: PS00513; ADENYLOSUCCLIN\_SYN\_2; 1.  
DR PROSITE: PS01266; ADENYLOSUCCLIN\_SYN\_1; 1.  
DR PFAM: PF00709; Adenylsucc\_synt; 1.  
DR HSP: P12283; LADI.  
KW PURINE BIOSYNTHESIS; LIGASE; GTP-BINDING.  
FT NP\_BIND 40 46 GTP (POTENTIAL).  
FT ACT\_SITE 172 172 BY SIMILARITY.  
FT ACT\_SITE 179 179 BY SIMILARITY.  
SQ SEQUENCE 457 AA; 50161 MW; 48EA29FB CRC32;  
Query Match 20.7%; Score 6; DB 1; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 4 VTSKTN 9  
QY 3 VTSKTN 8  
RESULT 9  
ID GAG\_EIAVY STANDARD: PRT; 486 AA.  
AC P03351;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEINS P15, P26, P11, P9].  
OS EQUINE INFECTIOUS ANEMIA VIRUS (ISOLATE WYOMING) (EIAV), AND  
OS EQUINE INFECTIOUS ANEMIA VIRUS (CLONE CL22) (EIAV).  
OC VIRUSES; RETROVIRUSES; RETROVIRIDAE; LENTIVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE WYOMING;  
RX MEDLINE: 86122873.  
RA STEPHENS R.M., CASEY J.W., RICE N.R.;  
RT "Equine infectious anemia virus gag and pol genes: relatedness to  
RT visna and AIDS virus.";  
RL SCIENCE 231:589-594(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CLONE CL22;  
RX MEDLINE: 92292230.

RA PERRY S.T., FLAHERTY M.T., KELLEY M.J., CLABOUGH D.L., TRONICK S.R.,  
RA COGGINS L., WHEAT L., LENGEL C.R., FULLER F.;  
RT "The surface envelope protein gene region of equine infectious anemia  
RT virus is not an important determinant of tropism in vitro.";  
RL J. VIROL. 66:4085-4097(1992).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M16575; G323837; -;  
DR EMBL; M87581; G290628; -;  
DR PIR; A03949; FOLJEV.  
DR PIR; A41991; FOLJ22.  
DR PFAM; PF00098; zf-CCHC; 2.  
DR PFAM; PF00607; gag\_p24; 1.  
DR HSP; P05888; 1AAF; 1.  
KW CORE PROTEIN; POLYPROTEIN.  
FT CHAIN 1 124  
FT CHAIN 125 359  
FT CHAIN 360 435  
FT CHAIN 436 486  
SQ SEQUENCE 486 AA; 54808 MW; 972A901C CRC32;

Query Match 20.7%; Score 6; DB 1; Length 486;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 276 NIROGA 281  
Qy 23 NIROGA 28

RESULT 10  
ID RAEP YEAST STANDARD; PRT; 603 AA.  
AC P32864;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE RAB PROTEINS GERANYLGERANYLTRANSFERASE COMPONENT A (RAB ESCORT  
DE PROTEIN) (REP).  
GN MS14 OR MRS6 OR YOR370C.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94179341.  
RA FUJIMURA K., TANAKA K., NAKANO A., TOH-E A.;  
RT "The Saccharomyces cerevisiae MS14 gene encodes the yeast counterpart  
RT of component A of Rab geranylgeranyltransferase.";  
RL J. BIOL. CHEM. 269:9205-9212(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 95188267.  
RA RAGNINI A., TEPLY R., WALDHERR M., VOSKOVA A., SCHWEYEN R.J.;  
RT "The yeast protein MRS6p, a homologue of the rabGDI and human  
RT chorioideraemia proteins, affects cytoplasmic and mitochondrial  
RT functions.";  
RL CURR. GENET. 26:308-314(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA DELIUS H., HEBLING U., HOFMANN B.;  
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [4]  
RP SEQUENCE OF 138-594 FROM N.A.  
RA WALDHERR M., VOSKOVA A., SCHWEYEN R.J.;  
RL SUBMITTED (APR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [5]

RP SIMILARITY TO CHOROIDEAEMIA PROTEIN.  
RX MEDLINE; 93251034.  
RA WALDHERR M., RAGNINI R.J., SCHWEYEN R.J., BOGUSKI M.S.;  
RT "MRS6 -- yeast homologue of the chorioideraemia gene.";  
RL NAT. GENET. 3:193-194(1993).  
CC -!- FUNCTION: BINDS UNPRENYLATED RAB, PRESENTS IT TO THE CATALYTIC  
CC COMPONENT B, AND REMAINS BOUND TO IT AFTER THE GERANYLGERANYL  
CC TRANSFER REACTION. THE COMPONENT A MAY BE REGENERATED BY  
CC TRANSFERRING ITS PRENYLATED RAB TO A PROTEIN ACCEPTOR.  
CC -!- SIMILARITY: BELONGS TO THE TCD/MRS6 FAMILY OF GDP DISSOCIATION  
CC INHIBITOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D26441; G452241; -;  
DR EMBL; X70339; G473151; -;  
DR EMBL; Z75278; E252194; -;  
DR EMBL; M90844; G171990; -;  
DR PIR; S27431; S27431.  
DR PIR; S47917; S47917.  
DR SGD; L0001194; MS14.  
DR PFAM; PF00996; GDI; 1.  
KW GTPASE ACTIVATION.  
FT CONFLICT 132 137  
FT CONFLICT 300 300  
SQ SEQUENCE 603 AA; 67374 MW; 43798BC8 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 603;  
Best Local Similarity 100.0%; Pred. No. 2.87e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 PSPLVW 29  
Qy 15 PSPLVW 20

RESULT 11  
ID GSPD KLEPN STANDARD; PRT; 660 AA.  
AC P15644;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE GENERAL SECRETION PATHWAY PROTEIN D PRECURSOR (PULLULANASE SECRETION  
DE ENVELOPE PULD).  
GN PULD.  
OS KLEBSIELLA PNEUMONIAE.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC KLEBSIELLA.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-42.  
RX MEDLINE; 90008916.  
RA D'ENFER C., REISS I., WANDERSMAN C., PUGSLEY A.P.;  
RT "Protein secretion by gram-negative bacteria. Characterization of two  
RT membrane proteins required for pullulanase secretion by Escherichia  
RT coli K-12.";  
RL J. BIOL. CHEM. 264:17462-17468(1989).  
CC -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE  
CC EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF PULLULANASE.  
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE EXED/OUTD/PULD/XPSD FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32613; G149303; -.
DR PIR; B31394; B31394.
DR PIR; B34469; B34469.
DR PROSITE; PS00875; T2SP.D: 1.
DR PFAM; PF00263; Bac_GSPDroteins; 1.
KW TRANSPORT; OUTER MEMBRANE; SIGNAL.
FT SIGNAL 1
FT CHAIN 28 660
SQ SEQUENCE 660 AA; 70658 MW; 3B7755D CRC32;

Query Match 20.7%; Score 6; DB 1; Length 660;
Best Local Similarity 100.0%; Pred. No. 2.87e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 215 NKDTSK 220
QY 8 NKDTSK 13

RESULT 12
ID VGLB_HSV7J STANDARD; PRT; 822 AA.
AC P52352.
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GLYCOPROTEIN B PRECURSOR.
GN GB OR U39.
OS HERPES SIMPLEX VIRUS (TYPE 7 / STRAIN J1) (HHV7).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA NICHOLAS J.;
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN B.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U43400; G1139641; -.
DR PFAM; PF00606; Glycoprotein_B; 1.
KW GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL.
FT SIGNAL 1
FT CHAIN 23 822
FT DOMAIN 23 822 GLYCOPROTEIN B.
FT TRANSMEM 23 649 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 650 670 POTENTIAL.
FT TRANSMEM 671 683 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 684 704 POTENTIAL.
FT DOMAIN 705 822 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 244 244 POTENTIAL.
FT CARBOHYD 324 324 POTENTIAL.
FT CARBOHYD 338 338 POTENTIAL.
FT CARBOHYD 352 352 POTENTIAL.
FT CARBOHYD 358 358 POTENTIAL.
FT CARBOHYD 387 387 POTENTIAL.
FT CARBOHYD 390 390 POTENTIAL.
FT CARBOHYD 482 482 POTENTIAL.
FT CARBOHYD 517 517 POTENTIAL.
FT CARBOHYD 567 567 POTENTIAL.
SQ SEQUENCE 822 AA; 93147 MW; 60FA195A CRC32;

Query Match 20.7%; Score 6; DB 1; Length 822;
Best Local Similarity 100.0%; Pred. No. 2.87e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1408 KFPSP 1413
QY 13 KFPSP 18

Query Match 20.7%; Score 6; DB 1; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2.87e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1408 KFPSP 1413
QY 13 KFPSP 18
```

```
Db 752 KTKDT 757
QY 6 KTKDT 11

RESULT 13
ID FAS2_PENPA STANDARD; PRT; 1857 AA.
AC P15368;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE FATTY ACID SYNTHASE, SUBUNIT ALPHA (EC 2.3.1.86) [CONTAINS:
DE EC 1.1.1.100; EC 2.3.1.41].
GN FAS2.
OS PENICILLIUM PATULUM (PENICILLIUM GRISEOFULVUM).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PLECTOMYCETES;
OC EUROTIALES; TRICHOCOMACEAE; MITOSPORIC TRICHOCOMACEAE; PENICILLIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89030697.
RA WIESNER P., BECK J., BECK K.-F., RPKA S., MUELLER G., LUECKE S.,
RA SCHWEIZER E.;
RT "Isolation and sequence analysis of the fatty acid synthetase FAS2
RT gene from Penicillium patulum";
RL EUR. J. BIOCHEM. 177:69-79(1988).
CC -!- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
CC THE ALPHA SUBUNIT CONTAINS DOMAINS FOR: ACYL CARRIER PROTEIN,
CC 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND 3-OXOACYL-[ACYL-
CC CARRIER-PROTEIN] SYNTHASE.
CC -!- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH -
CC LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADPH(+).
CC -!- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + MALONYL-[ACYL-
CC CARRIER PROTEIN] = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + CO(2) +
CC [ACYL-CARRIER PROTEIN].
CC -!- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
CC NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.
CC -!- SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF
CC TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).
CC -!- SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM
CC OTHER FUNGI.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37461; G169180; -.
DR PIR; S01787; S01787.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 1.
DR PROSITE; PS00606; B_KETOACYL-SYNTHASE; 1.
DR PFAM; PF00106; adh_short; 1.
KW FATTY ACID BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME; OXIDOREDUCTASE;
KW TRANSFERASE; NADP; PHOSPHOPANTHETHEINE.
FT DOMAIN 1 ? ACYL CARRIER.
FT DOMAIN 1 ? BETA-KETOACYL REDUCTASE.
FT DOMAIN ? 1857 BETA-KETOACYL SYNTHASE.
FT BINDING 174 174 PHOSPHOPANTHETHEINE (BY SIMILARITY).
FT ACT_SITE 1275 1275 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
SQ SEQUENCE 1857 AA; 204465 MW; ABF79D90 CRC32;

Query Match 20.7%; Score 6; DB 1; Length 1857;
Best Local Similarity 100.0%; Pred. No. 2.87e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 14
ID PSBF_EUGGR STANDARD; PRT; 41 AA.
AC P05334;
DT 01-NOV-1988 (REL. 09, CREATED)
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DE 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE CYTOCHROME B559 BETA CHAIN.
GN PSBF.
OS EUGLENA GRACILIS.
OG CHLOROPLAST.
OC EUKARYOTA; EUGLENOZOA; EUGLENIDA; EUGLENALES; EUGLENA.
RN [1]
RC STRAIN=Z;
RX MEDLINE; 88223485.
RA CUSHMAN J.C., CHRISTOPHER D.A., LITTLE M.C., HALLICK R.B., PRICE C.A.;
RT "Organization of the psbe, psbf, orf38, and orf42 gene loci on the
RT Euglena gracilis chloroplast genome.";
RL CURR. GENET. 13:173-180(1988).
CC -!- FUNCTION: THIS B-TYPE CYTOCHROME IS TIGHTLY ASSOCIATED WITH THE
CC REACTION CENTER OF PHOTOSYSTEM II AND POSSIBLY IS PART OF THE
CC WATER-OXIDATION COMPLEX.
CC -!- SUBUNIT: DIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SIMILARITY: BELONGS TO THE PSBE / PSBF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11874; G14363; -.
DR EMBL; X07073; G11494; -.
DR EMBL; X70810; G415752; -.
DR PIR; S00690; S00690.
DR PIR; S34517; S34517.
DR PROSITE; PS00537; CYTOCHROME_B559; 1.
DR PFAM; PF00283; cytochr_b559; 1.
DR MENDEL; 4491; EUGgr:psbf; 1.
DR CHLOROPLAST; PHOTOSYSTEM II;
FT DOMAIN 1 15 HEME; ELECTRON TRANSPORT; TRANSMEMBRANE.
FT TRANSMEM 16 40 POTENTIAL.
FT METAL 20 20 IRON (HEME) (BY SIMILARITY).
SQ SEQUENCE 41 AA; 4791 MW; 6299409C CRC32;

Query Match 17.2%; Score 5; DB 1; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.25e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 TNKDT 7
|
|
|
|
Qy 7 TNKDT 11

RESULT 15
ID PTHP_HAEIN STANDARD; PRT; 85 AA.
AC P43921;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PHOSPHOCARRIER PROTEIN HPR (HISTIDINE-CONTAINING PROTEIN).
GN PISH OR H11713.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=RD / KW20;
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
```

```
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLUM E., COTTON M.D.,
RA UTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE PHOSPHORYL GROUP FROM PHOSPHOENOLPYRUVATE
CC (PEP) IS TRANSFERRED TO THE PHOSPHORYL CARRIER PROTEIN HPR BY
CC ENZYME I. PHOSPHO-HPR THEN TRANSFERS IT TO THE PERMEASE (ENZYMES
CC II/III). HPR IS COMMON TO ALL PTS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: TO ALL OTHER HPR OR HPR DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32844; G1574568; -.
DR TIGR; H11713; -.
DR PROSITE; PS00369; PTS_HPR_HIS; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
DR PFAM; PF00381; PTS-HPR; 1.
DR HSP; P07008; 1PFH.
DR PHOSPHOTRANSFERASE SYSTEM; SUGAR TRANSPORT; PHOSPHORYLATION.
FT MOD_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 85 AA; 9004 MW; CF5DB671 CRC32;

Query Match 17.2%; Score 5; DB 1; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.25e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 ITVTS 37
|
|
|
|
Qy 1 ITVTS 5

Search completed: Sat Aug 28 14:56:41 1999
Job time : 44 secs.
```

\*\*\*\*\*  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:56:58 1999; MasPar time 7.59 Seconds  
Tabular output not generated.  
208.617 Million cell updates/sec  
Title: >US-09-049-696-46  
Description: (1-29) from US09049696.pep  
Perfect Score: 29  
Sequence: 1 ITVTSKNTKDTSEKPSPLVYVNIROGAS 29  
Scoring table: TABLE unitprotatable  
Gap 60  
Searched: 179066 seqs, 54579741 residues  
Post-processing: Minimum Match 0%  
Listing first 1000 summaries  
Database: sptrembl9  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus  
Statistics: Mean 2.722; Variance 0.392; scale 6.950  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
SUMMARIES

Result No.	Score	Query %	ID	Description	Pred. No.
1	6	20.7	5	COSMID F45E12.	5.03e+00
2	6	20.7	152	H03_ORF152 PROTEIN.	5.03e+00
3	6	20.7	205	ACETYL / ACYL TRANSFER	5.03e+00
4	6	20.7	225	ORF225.	5.03e+00
5	6	20.7	249	HYPOTHETICAL 27.2 KD P	5.03e+00
6	6	20.7	304	HYPOTHETICAL 32.8 KD P	5.03e+00
7	6	20.7	305	ORF 60.	5.03e+00
8	6	20.7	321	CYTCHROME AA3 QUINOL	5.03e+00
9	6	20.7	342	PROTEASOME REGULATORY	5.03e+00
10	6	20.7	363	C07A4.3 PROTEIN.	5.03e+00
11	6	20.7	365	GAG PROTEIN (FRAGMENT)	5.03e+00
12	6	20.7	385	CELL DIVISION PROTEIN.	5.03e+00
13	6	20.7	498	HYPOTHETICAL 54.5 KD P	5.03e+00
14	6	20.7	512	(GAG) GENE (POL.) (FRAG	5.03e+00
15	6	20.7	520	CONOPRESSIN RECEPTOR 2	5.03e+00
16	6	20.7	542	C44C10.4 PROTEIN.	5.03e+00
17	6	20.7	566	4-COUMARATE-COA LIGASE	5.03e+00
18	6	20.7	575	PILUS ASSEMBLY PROTEIN	5.03e+00
19	6	20.7	583	F411.7 PROTEIN.	5.03e+00
20	6	20.7	594	MRS6 PROTEIN.	5.03e+00

T05A1.3 PROTEIN.  
GPI60.  
GLOVEPROTEIN B.  
ENVELOPE GLYCOPROTEIN  
GLYCOPROTEIN B.  
F33E2.6 PROTEIN.  
GOB-5 PROTEIN.  
ORF 73.  
F56H1.3 PROTEIN.  
A19\_ORF1140 PROTEIN.  
ORF73 HOMOLOG.  
RIGUI.  
F58H7.6 PROTEIN.  
ANTIGEN CONTAINING EPI  
FATTY ACID SYNTHASE, A  
HYPOTHETICAL 215.3 KD  
XIN.  
ENVELOPE GLYCOPROTEIN,  
VIRILIS MITOCHONDRIAL,  
SURFACE ANTIGEN (FRAGM  
SURFACE ANTIGEN (FRAGM  
IMP (UL100) (FRAGMENT)  
IMP (UL100) (FRAGMENT)  
IMP (UL100) (FRAGMENT)  
MONOCARBOXYLATE TRANSP  
MHC CLASS IIB (FRAGMEN  
CATALASE (FRAGMENT).  
HYPOTHETICAL 10.5 KD P  
ENVELOPE GLYCOPROTEIN  
ENVELOPE GLYCOPROTEIN  
COAT PROTEIN (FRAGMENT  
COAT PROTEIN (FRAGMENT  
COAT PROTEIN (FRAGMENT  
COAT PROTEIN (FRAGMENT  
COAT PROTEIN (FRAGMENT  
HYPOTHETICAL 10.4 KD P  
CONSERVED HYPOTHETICAL  
RHEUMATOID FACTOR RF-E  
ENVELOPE GLYCOPROTEIN  
L71-9.  
103AA LONG HYPOTHETICA  
COMPLETE GENOME.  
LECTIN (FRAGMENT).  
LECTIN  
110AA LONG HYPOTHETICA  
ORF2 (FRAGMENT).  
LECTIN PRECURSOR.  
LECTIN.  
P450L1A1 GENE INTERNAL  
GP105 (FRAGMENT).  
SIMILAR TO.  
LECTIN.  
A\_IG002P16.14.  
CU/ZN SUPEROXIDE DISMU  
GENOME. PARTIAL SEQUEN  
WNT-1 PROTEIN (FRAGMEN  
HYPOTHETICAL 14.2 KD P  
ESTROGEN RECEPTOR (FRA  
GLUTAMATE-GATED CHLORI  
WUGSC:H\_RG013N12.GW.54  
L1 (FRAGMENT).  
HYPOTHETICAL 16.4 KD P  
L0045.  
RORF3.  
T15F16.12 PROTEIN.  
PUTATIVE ORF (FRAGMENT  
PAO SUBSTRAIN OT584 PY  
NADH DEHYDROGENASE SUB  
F38A1.9 PROTEIN.  
MHC CLASS IB (FRAGMENT

683 5 Q22187  
750 14 Q74105  
822 14 Q96910  
822 14 Q56280  
822 14 Q39989  
846 5 Q01699  
913 11 Q88826  
1089 14 Q40947  
1101 5 Q76369  
1140 2 P75405  
1162 14 Q98148  
1291 11 Q35973  
1419 5 Q45092  
1567 11 Q35243  
1858 3 P78615  
1969 3 Q14029  
2562 13 Q91957  
34 14 Q70340  
56 8 Q34398  
65 14 Q81104  
65 14 Q81103  
70 14 Q69190  
70 14 Q69193  
70 14 Q69204  
71 11 Q35440  
74 7 Q31518  
83 10 Q39634  
91 3 Q03788  
91 14 Q57052  
91 14 Q57048  
92 14 Q09733  
92 14 Q09732  
92 14 Q09735  
92 14 Q09731  
92 14 Q09734  
92 14 Q09730  
93 10 Q99198  
94 2 Q06623  
97 2 Q25361  
97 4 Q43234  
98 14 Q57056  
102 5 Q27252  
103 1 Q57777  
104 14 Q89496  
109 10 Q39988  
109 10 Q39991  
110 1 Q58002  
114 14 Q86910  
115 10 Q39987  
115 10 Q39989  
116 3 Q00585  
122 14 Q11912  
122 2 Q77336  
123 10 Q40422  
123 10 Q04644  
126 10 Q23785  
126 14 Q84412  
127 13 Q91229  
131 2 P73579  
133 6 Q28231  
141 5 Q23956  
142 4 Q60367  
148 14 Q11401  
149 10 Q49442  
152 2 Q85639  
152 2 Q52131  
153 10 Q81468  
153 4 Q14198  
155 2 P72118  
162 8 Q07212  
162 5 Q02256  
162 7 Q31513

21 6 20.7  
22 6 20.7  
23 6 20.7  
24 6 20.7  
25 6 20.7  
26 6 20.7  
27 6 20.7  
28 6 20.7  
29 6 20.7  
30 6 20.7  
31 6 20.7  
32 6 20.7  
33 6 20.7  
34 6 20.7  
35 6 20.7  
36 6 20.7  
37 6 20.7  
38 5 17.2  
39 5 17.2  
40 5 17.2  
41 5 17.2  
42 5 17.2  
43 5 17.2  
44 5 17.2  
45 5 17.2  
46 5 17.2  
47 5 17.2  
48 5 17.2  
49 5 17.2  
50 5 17.2  
51 5 17.2  
52 5 17.2  
53 5 17.2  
54 5 17.2  
55 5 17.2  
56 5 17.2  
57 5 17.2  
58 5 17.2  
59 5 17.2  
60 5 17.2  
61 5 17.2  
62 5 17.2  
63 5 17.2  
64 5 17.2  
65 5 17.2  
66 5 17.2  
67 5 17.2  
68 5 17.2  
69 5 17.2  
70 5 17.2  
71 5 17.2  
72 5 17.2  
73 5 17.2  
74 5 17.2  
75 5 17.2  
76 5 17.2  
77 5 17.2  
78 5 17.2  
79 5 17.2  
80 5 17.2  
81 5 17.2  
82 5 17.2  
83 5 17.2  
84 5 17.2  
85 5 17.2  
86 5 17.2  
87 5 17.2  
88 5 17.2  
89 5 17.2  
90 5 17.2  
91 5 17.2  
92 5 17.2  
93 5 17.2

94	5	17.2	163	2	Q53752	RIBONUCLEASE PRECURSOR	2.30e+02	167	5	17.2	256	14	Q96706	NUCLEAR SHUTTLE MOVEMENT	2.30e+02
95	5	17.2	163	2	O56849	MYF.	2.30e+02	168	5	17.2	256	5	P90657	IMMUNODOMINANT PIROPLA	2.30e+02
96	5	17.2	164	3	P78696	HIRSETULLIN A.	2.30e+02	169	5	17.2	256	5	O21585	SIMILAR TO HUMAN PLACE	2.30e+02
97	5	17.2	165	2	O48476	SIMILAR TO RHODOBACTER	2.30e+02	170	5	17.2	257	5	O61161	MAJOR PIROPLASM SURFAC	2.30e+02
98	5	17.2	165	2	O34612	HYPOTHETICAL 18.3 KD P	2.30e+02	171	5	17.2	258	2	P73671	HYDROGENASE ACCESSORY	2.30e+02
99	5	17.2	175	5	O17265	BPCDNA6 PROTEIN.	2.30e+02	172	5	17.2	259	2	Q51125	HYPOXANTHINE GEM	2.30e+02
100	5	17.2	178	2	P81380	UBIQUINOL-CYTOCHROME C	2.30e+02	173	5	17.2	260	5	O76879	HYPOXANTHINE 29.2 KD P	2.30e+02
101	5	17.2	178	2	P96729	HYPOTHETICAL 19.1 KD P	2.30e+02	174	5	17.2	260	5	O24764	0.9-KB RNA TRANSCRIPT.	2.30e+02
102	5	17.2	178	2	O48961	SIMILAR TO GTP-BIND. G	2.30e+02	175	5	17.2	260	5	O80861	F7E1.17 PROTEIN.	2.30e+02
103	5	17.2	181	2	O53573	ORF C.	2.30e+02	176	5	17.2	260	10	O80861	MYOSIN II NONMUSCLE (F	2.30e+02
104	5	17.2	181	2	O87052	248F PROTEIN (FRAGMENT	2.30e+02	177	5	17.2	261	13	Q91300	MYOSIN II NONMUSCLE (F	2.30e+02
105	5	17.2	182	14	O90057	THYMIDINE KINASE.	2.30e+02	178	5	17.2	261	13	Q91301	MYOSIN II NONMUSCLE (F	2.30e+02
106	5	17.2	182	14	O69111	(TK- STRAIN 333) THYM	2.30e+02	179	5	17.2	261	5	O61160	MAJOR PIROPLASM SURFAC	2.30e+02
107	5	17.2	182	2	P96401	HYPOTHETICAL 20.1 KD P	2.30e+02	180	5	17.2	262	10	O43630	ALPHA-AMYLASE INHIBIT	2.30e+02
108	5	17.2	185	2	O86434	PUTATIVE ACETYL TRANSF	2.30e+02	181	5	17.2	265	14	P89200	VIRUS COAT PROTEIN (FR	2.30e+02
109	5	17.2	187	2	O34872	YTB.	2.30e+02	182	5	17.2	265	14	P90374	VIRUS COAT PROTEIN (FR	2.30e+02
110	5	17.2	188	2	P74118	HYPOTHETICAL 21.0 KD P	2.30e+02	183	5	17.2	265	10	Q01660	ZEIN-ALPHA PRECURSOR (	2.30e+02
111	5	17.2	189	2	O53086	LPHA PROTEIN (DOTK).	2.30e+02	184	5	17.2	266	10	O48969	22-KDA ALPHA ZEIN 8.	2.30e+02
112	5	17.2	190	2	O52586	HYPOTHETICAL 20.4 KD P	2.30e+02	185	5	17.2	266	14	O88190	COAT PROTEIN.	2.30e+02
113	5	17.2	191	14	O36169	21.6K PROTEIN.	2.30e+02	186	5	17.2	267	14	O88196	COAT PROTEIN (FRAGMENT	2.30e+02
114	5	17.2	192	14	O89367	HYPOTHETICAL PROTEIN.	2.30e+02	187	5	17.2	267	10	P93019	BRASSICA NAPUS HYPOTHE	2.30e+02
115	5	17.2	200	10	O40741	LEA-LIKE PROTEIN.	2.30e+02	188	5	17.2	268	5	O00848	MAJOR PIROPLASM SURFAC	2.30e+02
116	5	17.2	200	10	O40696	GROUP 3 LEA (TYPE I) P	2.30e+02	189	5	17.2	268	13	Q91305	MYOSIN VII (FRAGMENT).	2.30e+02
117	5	17.2	203	2	O32811	X30 (X30) (FRAGMENT).	2.30e+02	190	5	17.2	268	4	O99618	CHROMOSOME 12P13 SEQUE	2.30e+02
118	5	17.2	204	10	O80329	HEAT SHOCK PROTEIN 26	2.30e+02	191	5	17.2	269	5	O16070	82 KDA HEAT SHOCK PROT	2.30e+02
119	5	17.2	205	10	O80328	HEAT SHOCK PROTEIN 26	2.30e+02	192	5	17.2	269	5	O16078	82 KDA HEAT SHOCK PROT	2.30e+02
120	5	17.2	206	10	O48787	PUTATIVE BLUE COPPER P	2.30e+02	193	5	17.2	269	5	O16084	82 KDA HEAT SHOCK PROT	2.30e+02
121	5	17.2	207	11	O63580	TITE (FRAGMENT).	2.30e+02	194	5	17.2	269	5	O16080	82 KDA HEAT SHOCK PROT	2.30e+02
122	5	17.2	209	2	O06740	YITE PROTEIN.	2.30e+02	195	5	17.2	269	5	O16083	82 KDA HEAT SHOCK PROT	2.30e+02
123	5	17.2	210	2	P73288	HYPOTHETICAL 22.0 KD P	2.30e+02	196	5	17.2	269	5	O16074	82 KDA HEAT SHOCK PROT	2.30e+02
124	5	17.2	211	5	O45522	F45B8.3 PROTEIN.	2.30e+02	197	5	17.2	269	5	O16080	82 KDA HEAT SHOCK PROT	2.30e+02
125	5	17.2	218	14	O84752	FMDV GENES FOR THE STR	2.30e+02	198	5	17.2	269	5	O16082	82 KDA HEAT SHOCK PROT	2.30e+02
126	5	17.2	218	14	O84762	FMDV RNA OF PRIMARY TR	2.30e+02	199	5	17.2	269	5	O16081	82 KDA HEAT SHOCK PROT	2.30e+02
127	5	17.2	218	14	O84756	STRAIN C1 TRANSLATION	2.30e+02	200	5	17.2	269	5	O16079	82 KDA HEAT SHOCK PROT	2.30e+02
128	5	17.2	218	14	O65035	A10 (A-HOLLAND) CAPSID	2.30e+02	201	5	17.2	269	5	O16075	82 KDA HEAT SHOCK PROT	2.30e+02
129	5	17.2	219	2	O53625	MITHRAMYCIN RESISTANCE	2.30e+02	202	5	17.2	269	5	O16071	82 KDA HEAT SHOCK PROT	2.30e+02
130	5	17.2	219	5	O02210	C01A2.6 PROTEIN.	2.30e+02	203	5	17.2	269	5	O16072	82 KDA HEAT SHOCK PROT	2.30e+02
131	5	17.2	220	5	O17705	C06B3.6 PROTEIN.	2.30e+02	204	5	17.2	269	5	O16069	82 KDA HEAT SHOCK PROT	2.30e+02
132	5	17.2	221	10	O22944	UNKNOWN PROTEIN.	2.30e+02	205	5	17.2	269	5	O16087	82 KDA HEAT SHOCK PROT	2.30e+02
133	5	17.2	225	3	O42881	LYSOPHOSPHOLIPASE.	2.30e+02	206	5	17.2	269	5	O16073	82 KDA HEAT SHOCK PROT	2.30e+02
134	5	17.2	225	3	O74194	CYTOCHROME P450 LIAL D	2.30e+02	207	5	17.2	269	5	O16089	82 KDA HEAT SHOCK PROT	2.30e+02
135	5	17.2	226	14	O90663	SURFACE ANTIGEN.	2.30e+02	208	5	17.2	269	5	O16088	82 KDA HEAT SHOCK PROT	2.30e+02
136	5	17.2	226	5	O93830	F59C6.4 PROTEIN.	2.30e+02	209	5	17.2	269	5	O16085	82 KDA HEAT SHOCK PROT	2.30e+02
137	5	17.2	228	14	O55261	THYMIDINE KINASE.	2.30e+02	210	5	17.2	269	5	O00814	SERINE/THREONINE PROTE	2.30e+02
138	5	17.2	229	2	O53710	UNKNOWN ORF1.	2.30e+02	211	5	17.2	269	5	O16086	82 KDA HEAT SHOCK PROT	2.30e+02
139	5	17.2	229	14	O10689	TRIPLE GENE BLOCK 1.	2.30e+02	212	5	17.2	269	5	O16088	82 KDA HEAT SHOCK PROT	2.30e+02
140	5	17.2	230	14	O8454	MATRIX PROTEIN 1 (M1 P	2.30e+02	213	5	17.2	269	5	O16076	82 KDA HEAT SHOCK PROT	2.30e+02
141	5	17.2	230	14	O82681	POLYMERASE-ASSOCIATED	2.30e+02	214	5	17.2	269	2	P73131	HYPOTHETICAL 29.3 KD P	2.30e+02
142	5	17.2	233	14	O37154	MOVEMENT PROTEIN 1.	2.30e+02	215	5	17.2	269	5	O16077	82 KDA HEAT SHOCK PROT	2.30e+02
143	5	17.2	234	2	O07287	OPACITY PROTEIN (FRAGM	2.30e+02	216	5	17.2	269	5	O18670	HEAT SHOCK PROTEIN 83	2.30e+02
144	5	17.2	234	10	O03993	CHROMOPLAST-ASSOCIATED	2.30e+02	217	5	17.2	269	5	O61306	82 KDA HEAT SHOCK PROT	2.30e+02
145	5	17.2	234	5	O25769	ASPARAGINE-RICH ANTIGE	2.30e+02	218	5	17.2	269	10	Q93361	HYPOTHETICAL 31.8 KD P	2.30e+02
146	5	17.2	236	6	Q95197	ZINC FINGER PROTEIN 15	2.30e+02	219	5	17.2	270	2	O31917	YOOV PROTEIN.	2.30e+02
147	5	17.2	237	5	O93652	F30F8.5 PROTEIN.	2.30e+02	220	5	17.2	270	9	O64130	YOOV PROTEIN.	2.30e+02
148	5	17.2	240	2	P74888	GLUTAMINE TRANSPORT AT	2.30e+02	221	5	17.2	272	11	O62707	NONMUSCLE MYOSIN HEAVY	2.30e+02
149	5	17.2	240	10	O81961	HEAT SHOCK PROTEIN 26	2.30e+02	222	5	17.2	272	5	O19097	CODING FOR BY C. ELEGAN	2.30e+02
150	5	17.2	241	14	O84321	REPLICATION PROTEIN E1	2.30e+02	223	5	17.2	272	2	O44376	OPHE.	2.30e+02
151	5	17.2	241	5	O21577	COSMID M6.	2.30e+02	224	5	17.2	273	2	O34563	GLUTAMINE ABC TRANSPOR	2.30e+02
152	5	17.2	244	5	O76621	Y57G7A.5 PROTEIN.	2.30e+02	225	5	17.2	273	5	O01434	COSMID B0207.	2.30e+02
153	5	17.2	245	2	O67916	HYPOTHETICAL 26.9 KD P	2.30e+02	226	5	17.2	275	2	O50261	AGAC.	2.30e+02
154	5	17.2	245	5	O76847	HOMEOBOX PROTEIN (FRAG	2.30e+02	227	5	17.2	276	1	O29787	CONSERVED HYPOTHETICAL	2.30e+02
155	5	17.2	249	7	O31514	MHC CLASS IIB.	2.30e+02	228	5	17.2	276	10	O24482	SALIS-2.	2.30e+02
156	5	17.2	250	8	O02673	HYPOTHETICAL 29.3 KD P	2.30e+02	229	5	17.2	278	14	P89139	LARGE GLYCOPROTEIN.	2.30e+02
157	5	17.2	251	3	O12592	3-ISOPROPYLMALATE DEHY	2.30e+02	230	5	17.2	278	14	Q90092	POLYPROTEIN 3BCD.	2.30e+02
158	5	17.2	252	1	O58345	252AA LONG HYPOTHETICA	2.30e+02	231	5	17.2	279	5	P90659	32 KDA IMMUNODOMINANT	2.30e+02
159	5	17.2	253	3	O42737	NECROSIS-INDUCING PEPT	2.30e+02	232	5	17.2	279	3	Q12164	CHROMOSOME XII READING	2.30e+02
160	5	17.2	253	14	O67583	BEAN GOLDEN MOSAIC GEM	2.30e+02	233	5	17.2	279	5	O00944	32 KDA IMMUNODOMINANT	2.30e+02
161	5	17.2	254	14	O66215	MOVEMENT PROTEIN.	2.30e+02	234	5	17.2	280	10	O43461	NITRATE REDUCTASE (FRA	2.30e+02
162	5	17.2	254	14	O41944	HYPOTHETICAL 29.0 KD P	2.30e+02	235	5	17.2	280	2	O53316	S-ADENOSYL METHIONINE-	2.30e+02
163	5	17.2	255	2	O50731	STAGE 0 SPORULATION PR	2.30e+02	236	5	17.2	281	1	O57840	281AA LONG HYPOTHETICA	2.30e+02
164	5	17.2	256	2	O51124	OPACITY OUTERMEMBRANE	2.30e+02	237	5	17.2	281	5	O16494	B0238.3 PROTEIN.	2.30e+02
165	5	17.2	256	14	O84397	DNA-A COMPONENT.	2.30e+02	238	5	17.2	282	11	O62706	NONMUSCLE MYOSIN HEAVY	2.30e+02
166	5	17.2	256	5	O00943	32 KDA IMMUNODOMINANT	2.30e+02	239	5	17.2	283	5	O27401	TS-32K PROTEIN PRECURS	2.30e+02

240	5	17.2	283	5	026792	PIROPLASM MAJOR IMMUNO	2.30e+02	313	5	17.2	329	14	090253	POLYPROTEIN (FRAGMENT)	2.30e+02
241	5	17.2	283	5	076176	PIROPLASM MAJOR IMMUNO	2.30e+02	314	5	17.2	329	14	057011	POLYPROTEIN (FRAGMENT)	2.30e+02
242	5	17.2	283	5	027077	PIROPLASM MAJOR IMMUNO	2.30e+02	315	5	17.2	329	14	057009	POLYPROTEIN (FRAGMENT)	2.30e+02
243	5	17.2	283	5	015928	PIROPLASM MAJOR IMMUNO	2.30e+02	316	5	17.2	330	14	057007	POLYPROTEIN (FRAGMENT)	2.30e+02
244	5	17.2	283	5	076175	PIROPLASM MAJOR IMMUNO	2.30e+02	317	5	17.2	330	14	096668	COMPEA APHID BORNE MOS	2.30e+02
245	5	17.2	283	14	089251	COAT PROTEIN.	2.30e+02	318	5	17.2	331	13	073667	MELANOCORTIN 4-RECEPT	2.30e+02
246	5	17.2	284	5	076174	PIROPLASM MAJOR IMMUNO	2.30e+02	319	5	17.2	332	2	059830	CLASS A BETA-LACTAMASE	2.30e+02
247	5	17.2	284	3	078777	FISSION YEAST.	2.30e+02	320	5	17.2	332	2	059850	IMMUNOGENIC PROTEIN P3	2.30e+02
248	5	17.2	286	14	065403	COAT PROTEIN (FRAGMENT	2.30e+02	321	5	17.2	332	2	033981	IMMUNOGENIC PROTEIN P3	2.30e+02
249	5	17.2	286	14	065406	COAT PROTEIN (FRAGMENT	2.30e+02	322	5	17.2	332	2	033981	IMMUNOGENIC PROTEIN P3	2.30e+02
250	5	17.2	286	14	065430	COAT PROTEIN (FRAGMENT	2.30e+02	323	5	17.2	335	4	043403	RAD51-INTERACTING PROT	2.30e+02
251	5	17.2	287	9	038585	ANTI-REPRESSOR PROTEIN	2.30e+02	324	5	17.2	335	2	086132	ORNITHINE CARBAMOYLTRA	2.30e+02
252	5	17.2	287	14	065401	COAT PROTEIN (FRAGMENT	2.30e+02	325	5	17.2	336	6	018906	PLATELET AND T CELL AC	2.30e+02
253	5	17.2	287	14	065424	COAT PROTEIN (FRAGMENT	2.30e+02	326	5	17.2	336	6	046551	PLATELET AND T CELL AC	2.30e+02
254	5	17.2	287	14	065424	COAT PROTEIN (FRAGMENT	2.30e+02	327	5	17.2	336	1	030070	FLAVOPROTEIN (FPRA-1).	2.30e+02
255	5	17.2	287	5	021125	KO2B9.3 PROTEIN.	2.30e+02	328	5	17.2	336	4	015762	DYMAT-1.	2.30e+02
256	5	17.2	287	14	085076	CAPSID PROTEIN (FRAGME	2.30e+02	329	5	17.2	336	2	069207	HYPOTHETICAL 35.7 KD P	2.30e+02
257	5	17.2	287	14	065428	COAT PROTEIN (FRAGMENT	2.30e+02	330	5	17.2	336	4	016096	LYMPHOCYTE-SPECIFIC PR	2.30e+02
258	5	17.2	287	14	065404	COAT PROTEIN (FRAGMENT	2.30e+02	331	5	17.2	337	5	016493	B0238.4 PROTEIN.	2.30e+02
259	5	17.2	287	14	065429	COAT PROTEIN (FRAGMENT	2.30e+02	332	5	17.2	337	1	005169	3-ISOPROPYLMALATE DEHY	2.30e+02
260	5	17.2	287	14	065932	COAT PROTEIN (FRAGMENT	2.30e+02	333	5	17.2	339	14	090709	POLYPEPTIDE (FRAGMENT)	2.30e+02
261	5	17.2	288	4	013210	ADRENOLEUKODYSTROPHY R	2.30e+02	334	5	17.2	341	10	093714	PETHY, ZPT2-6.	2.30e+02
262	5	17.2	288	2	050848	IMMUNOGENIC PROTEIN P3	2.30e+02	335	5	17.2	341	2	054765	HYPOTHETICAL 37.5 KD P	2.30e+02
263	5	17.2	289	5	022133	T04AB.4 PROTEIN.	2.30e+02	336	5	17.2	342	14	011942	ENVELOPE GLYCOPROTEIN	2.30e+02
264	5	17.2	290	14	090710	POLYPEPTIDE (FRAGMENT)	2.30e+02	337	5	17.2	342	5	077363	MAL3P4.5B PROTEIN.	2.30e+02
265	5	17.2	290	5	002561	PIROPLASM SURFACE ANTI	2.30e+02	338	5	17.2	342	2	069767	P38K.	2.30e+02
266	5	17.2	293	14	065420	COAT PROTEIN (FRAGMENT	2.30e+02	339	5	17.2	343	2	055989	HYPOTHETICAL 38.4 KD P	2.30e+02
267	5	17.2	294	14	065423	COAT PROTEIN (FRAGMENT	2.30e+02	340	5	17.2	344	2	033323	OXIDOREDUCTASE.	2.30e+02
268	5	17.2	295	14	066408	COAT PROTEIN (FRAGMENT	2.30e+02	341	5	17.2	344	2	067238	PHOSPHO-N-ACETYLMURAMO	2.30e+02
269	5	17.2	296	14	092268	NIB PROTEIN/COAT PROTE	2.30e+02	342	5	17.2	344	2	075340	PHOSPHO-N-ACETYLMURAMO	2.30e+02
270	5	17.2	296	14	057010	RNA FOR NIB PROTEIN AN	2.30e+02	343	5	17.2	345	2	075340	MG307 HOMOLOG.	2.30e+02
271	5	17.2	299	1	058693	299AA LONG HYPOTHETICA	2.30e+02	344	5	17.2	347	2	051584	TRANSMEMBRANE GLYCOPRO	2.30e+02
272	5	17.2	299	1	058693	299AA LONG HYPOTHETICA	2.30e+02	345	5	17.2	348	2	051584	SPERMIDINE/PUTRESCINE	2.30e+02
273	5	17.2	299	14	064794	1-EVIDENCE-PREDICTED B	2.30e+02	346	5	17.2	348	5	017431	THIOL PROTEASE (FRAGME	2.30e+02
274	5	17.2	299	14	064079	POLYPROTEIN (FRAGMENT	2.30e+02	347	5	17.2	351	9	038350	GLUTAMATE-GATED CHLORI	2.30e+02
275	5	17.2	300	14	090229	POLYPROTEIN (FRAGMENT)	2.30e+02	348	5	17.2	351	5	025634	NIB PROTEIN/COAT PROTE	2.30e+02
276	5	17.2	300	14	090232	POLYPROTEIN (FRAGMENT)	2.30e+02	349	5	17.2	353	14	092269	VAA SURFACE LIPOPROTEI	2.30e+02
277	5	17.2	300	14	090230	POLYPROTEIN (FRAGMENT)	2.30e+02	350	5	17.2	353	2	053101	TRANSCRIPTION FACTOR V	2.30e+02
278	5	17.2	300	14	090228	POLYPROTEIN (FRAGMENT)	2.30e+02	351	5	17.2	355	14	089919	POLYPROTEIN (FRAGMENT)	2.30e+02
279	5	17.2	300	14	090231	POLYPROTEIN (FRAGMENT)	2.30e+02	352	5	17.2	355	13	073679	POLYPROTEIN (FRAGMENT)	2.30e+02
280	5	17.2	302	3	005330	PUTATIVE ORF.	2.30e+02	353	5	17.2	357	2	055458	HYPOTHETICAL 38.4 KD P	2.30e+02
281	5	17.2	302	3	071853	HYPOTHETICAL 31.1 KD P	2.30e+02	354	5	17.2	362	3	059723	HYPOTHETICAL 41.3 KD P	2.30e+02
282	5	17.2	304	5	044694	C45H4.10 PROTEIN.	2.30e+02	355	5	17.2	365	2	031301	ELONGATION FACTOR EF-T	2.30e+02
283	5	17.2	305	14	066662	RIBONUCLEOTIDE REDUCTA	2.30e+02	356	5	17.2	365	14	089735	POLYPROTEIN (FRAGMENT)	2.30e+02
284	5	17.2	305	14	068997	COAT PROTEIN (FRAGMENT	2.30e+02	357	5	17.2	367	14	089736	POLYPROTEIN (FRAGMENT)	2.30e+02
285	5	17.2	306	14	065425	COAT PROTEIN (FRAGMENT	2.30e+02	358	5	17.2	367	14	089736	POLYPROTEIN (FRAGMENT)	2.30e+02
286	5	17.2	306	5	045386	F21H7.5 PROTEIN.	2.30e+02	359	5	17.2	370	13	096675	NIB PROTEIN/COAT PROTE	2.30e+02
287	5	17.2	307	1	028231	HYPOTHETICAL 35.1 KD P	2.30e+02	360	5	17.2	370	3	040083	WNT1.	2.30e+02
288	5	17.2	308	14	065426	COAT PROTEIN (FRAGMENT	2.30e+02	361	5	17.2	371	2	050503	D9461.24P.	2.30e+02
289	5	17.2	308	14	089369	SIMILAR TO SCHIZOSACCH	2.30e+02	362	5	17.2	371	2	050503	PERIPLASMIC SOLUTE-BIN	2.30e+02
290	5	17.2	309	14	009722	COAT PROTEIN PRECURSOR	2.30e+02	363	5	17.2	373	3	012591	3-ISOPROPYLMALATE DEHY	2.30e+02
291	5	17.2	309	14	090695	POLYPEPTIDE (FRAGMENT)	2.30e+02	364	5	17.2	373	3	012591	3-ISOPROPYLMALATE DEHY	2.30e+02
292	5	17.2	309	14	065427	POLYPROTEIN (FRAGMENT)	2.30e+02	365	5	17.2	374	4	080591	T2711.2 PROTEIN.	2.30e+02
293	5	17.2	309	14	065422	COAT PROTEIN (FRAGMENT)	2.30e+02	366	5	17.2	374	10	080591	THYMIDINE KINASE.	2.30e+02
294	5	17.2	309	6	028667	LOW-DENSITY LIPOPROTEI	2.30e+02	367	5	17.2	376	14	086630	THYMIDINE KINASE.	2.30e+02
295	5	17.2	310	2	008454	PECTATE LYASE (EC 4.2.	2.30e+02	368	5	17.2	376	14	086630	THYMIDINE KINASE.	2.30e+02
296	5	17.2	312	14	089122	RNA REPLICASE READ THR	2.30e+02	369	5	17.2	377	13	091997	3BETA-HYDROXYSTEROID D	2.30e+02
297	5	17.2	312	2	044420	SUGAR-BINDING TRANSPOR	2.30e+02	370	5	17.2	377	2	051302	HYPOTHETICAL 44.2 KD P	2.30e+02
298	5	17.2	313	10	080936	F13M22.18 PROTEIN.	2.30e+02	371	5	17.2	378	3	059925	POLYGALACTURONASE.	2.30e+02
299	5	17.2	313	5	022569	T19CA.2 PROTEIN.	2.30e+02	372	5	17.2	381	4	075584	MALIC ENZYME.	2.30e+02
300	5	17.2	314	14	086659	GENOME POLYPROTEIN [CO	2.30e+02	373	5	17.2	382	2	047844	ESTROGEN RECEPTOR BETA	2.30e+02
301	5	17.2	314	14	086656	GENOME POLYPROTEIN [CO	2.30e+02	374	5	17.2	382	2	047844	ESTROGEN RECEPTOR BETA	2.30e+02
302	5	17.2	314	14	086657	GENOME POLYPROTEIN [CO	2.30e+02	375	5	17.2	383	5	001838	KIAA0468 PROTEIN.	2.30e+02
303	5	17.2	314	2	073125	HYPOTHETICAL 34.1 KD P	2.30e+02	376	5	17.2	384	4	075056	PRE-S1, PRE-S2 AND S.	2.30e+02
304	5	17.2	315	2	073036	REGULATORY COMPONENTS	2.30e+02	377	5	17.2	385	14	039888	PRE-S1, PRE-S2 AND S.	2.30e+02
305	5	17.2	317	1	027997	CAROTENOID BIOSYNTHETI	2.30e+02	378	5	17.2	387	14	039887	PRE-S1, PRE-S2 AND S.	2.30e+02
306	5	17.2	318	10	080629	POTATIVE LYSOPHOSPHOLI	2.30e+02	379	5	17.2	388	2	053111	OLIGOPEPTIDE TRANSPORT	2.30e+02
307	5	17.2	319	2	052927	EXOW.	2.30e+02	380	5	17.2	388	1	058427	388AA LONG HYPOTHETICA	2.30e+02
308	5	17.2	325	5	020213	F40F12.3 PROTEIN.	2.30e+02	381	5	17.2	389	5	024977	SIMILAR TO STEROID/THY	2.30e+02
309	5	17.2	326	14	093129	POLYPROTEIN (FRAGMENT)	2.30e+02	382	5	17.2	391	2	024977	MOLYBDOPROTEIN BIOSYNTH	2.30e+02
310	5	17.2	327	10	064592	F1707.2.	2.30e+02	383	5	17.2	392	2	084345	HEAT SHOCK PROTEIN J.	2.30e+02
311	5	17.2	327	3	007379	CHROMOSOME IV READING	2.30e+02	384	5	17.2	392	14	086995	POLYPROTEIN PRECURSOR	2.30e+02
312	5	17.2	328	14	057008	POLYPROTEIN (FRAGMENT)	2.30e+02	385	5	17.2	397	11	035118	GLIAL CELL LINE DERIVE	2.30e+02
											397	11	055243	GLIAL CELL LINE-DERIVE	2.30e+02

386	5	17.2	397	11	035325	GLIAL CELL LINE-DERIVE	2.30e+02	459	5	17.2	493	2	067365	HYPOTHETICAL 56.5 KD P	2.30e+02
387	5	17.2	399	2	031663	YKRT PROTEIN.	2.30e+02	460	5	17.2	494	5	04275	VARIANT SURFACE GLYP	2.30e+02
388	5	17.2	400	2	052293	UXPB, UXPA, XCPP, XCPQ	2.30e+02	461	5	17.2	495	4	060608	ESTROGEN RECEPTOR BETA	2.30e+02
389	5	17.2	401	4	060609	GNF, FAMILY RECEPTOR A	2.30e+02	462	5	17.2	495	2	089080	RNA POLYMERASE SUBUNIT	2.30e+02
390	5	17.2	404	5	017183	F34D6.1 PROTEIN.	2.30e+02	463	5	17.2	498	10	042435	CAPSANTHIN/CAPSORUBIN	2.30e+02
391	5	17.2	403	10	081448	T27D20.7 PROTEIN.	2.30e+02	464	5	17.2	498	8	021287	NADH DEHYDROGENASE, SU	2.30e+02
392	5	17.2	404	5	091148	SIMILAR TO MOUSE TESTO	2.30e+02	465	5	17.2	498	10	023235	ACC SYNTHASE (EC 4.4.1	2.30e+02
393	5	17.2	405	14	088563	COAT PROTEIN (FRAGMENT	2.30e+02	466	5	17.2	500	4	P78427	MYOSIN VIIA (FRAGMENT)	2.30e+02
394	5	17.2	410	2	033331	DEHYDROGENASE.	2.30e+02	467	5	17.2	501	3	P74779	SERINE PROLINE RICH PR	2.30e+02
395	5	17.2	410	2	P72973	NA/H ANTIporter.	2.30e+02	468	5	17.2	502	5	045485	F38A1.5 PROTEIN.	2.30e+02
396	5	17.2	411	13	092031	METALLOPROTEINASE PREC	2.30e+02	469	5	17.2	506	10	065727	SQUALENE EPOXIDASE HOM	2.30e+02
397	5	17.2	413	2	008340	GLUTAMATE SYNTHASE (EC	2.30e+02	470	5	17.2	506	5	021518	SIMILARITY TO CLASS-II	2.30e+02
398	5	17.2	413	5	021545	M153.2 PROTEIN.	2.30e+02	471	5	17.2	506	8	047162	RIBOSOMAL MATURASE.	2.30e+02
399	5	17.2	419	5	044446	CO2B10.1 PROTEIN.	2.30e+02	472	5	17.2	507	14	090233	POLYPROTEIN (FRAGMENT)	2.30e+02
400	5	17.2	420	4	013641	5T4 ONCOFETAL ANTIGEN	2.30e+02	473	5	17.2	508	11	070420	CORTACTIN ISOFORM B.	2.30e+02
401	5	17.2	420	2	087565	NATB.	2.30e+02	474	5	17.2	511	5	Q21252	KOGA4.1 PROTEIN.	2.30e+02
402	5	17.2	420	5	016285	F59A7.1 PROTEIN.	2.30e+02	475	5	17.2	512	13	Q90670	ACTIVIN RECEPTOR IIB.	2.30e+02
403	5	17.2	422	2	006321	HYPOTHETICAL 45.8 KD P	2.30e+02	476	5	17.2	512	6	095126	ACTIVIN RECEPTOR TYPE	2.30e+02
404	5	17.2	422	10	023764	CRATROSTIGMA PLANTIGI	2.30e+02	477	5	17.2	512	10	082415	TYROSINE/DOPA DECARBOX	2.30e+02
405	5	17.2	423	2	P76198	FROM BASES 1766777 TO	2.30e+02	478	5	17.2	513	4	060685	ESTROGEN RECEPTOR BETA	2.30e+02
406	5	17.2	424	5	016445	C54F6.11 PROTEIN.	2.30e+02	479	5	17.2	514	10	085404	SQUALENE EPOXIDASE HOM	2.30e+02
407	5	17.2	425	3	Q12741	CAMP-DEPENDENT PROTEIN	2.30e+02	480	5	17.2	515	11	034949	NEMO LIKE KINASE.	2.30e+02
408	5	17.2	427	14	069319	UL48H.	2.30e+02	481	5	17.2	517	3	Q12571	LACASE PRECURSOR (EC	2.30e+02
409	5	17.2	427	13	Q90603	MDV ALPHA TIF.	2.30e+02	482	5	17.2	517	3	Q07383	HISTIDINE KINASE.	2.30e+02
410	5	17.2	430	5	001982	F01D4.8 PROTEIN.	2.30e+02	483	5	17.2	517	10	065402	SQUALENE EPOXIDASE HOM	2.30e+02
411	5	17.2	432	3	014452	HYPOTHETICAL 61.4KD PR	2.30e+02	484	5	17.2	518	5	Q21528	M110.1 PROTEIN.	2.30e+02
412	5	17.2	433	5	044473	E04A4.3 PROTEIN.	2.30e+02	485	5	17.2	520	3	059944	LACASE PRECURSOR (EC	2.30e+02
413	5	17.2	434	5	P73276	SENSORY TRANSDUCTION H	2.30e+02	486	5	17.2	520	3	013421	PHENOLOXIDASE (EC 1.10	2.30e+02
414	5	17.2	435	5	019374	F13B12.4 PROTEIN.	2.30e+02	487	5	17.2	521	10	048639	INORGANIC PHOSPHATE TR	2.30e+02
415	5	17.2	438	5	046124	GBR-2B PROTEIN PRECURS	2.30e+02	488	5	17.2	521	10	043811	PUTATIVE PROTON/PHOSPH	2.30e+02
416	5	17.2	440	2	034936	YTCQ.	2.30e+02	489	5	17.2	522	2	059929	CHITINASE PRECURSOR (E	2.30e+02
417	5	17.2	440	2	046216	VIRS.	2.30e+02	490	5	17.2	523	14	P89903	FROM AFRICAN GREEN MON	2.30e+02
418	5	17.2	442	14	056833	POLYPROTEIN (FRAGMENTS	2.30e+02	491	5	17.2	523	4	Q13080	TPR-MET FUSION PROTEIN	2.30e+02
419	5	17.2	448	10	081452	T27D20.3 PROTEIN.	2.30e+02	492	5	17.2	525	10	082417	TYROSINE/DOPA DECARBOX	2.30e+02
420	5	17.2	448	4	075021	MONOCYTE IMMUNOGLOBUL	2.30e+02	493	5	17.2	526	3	013456	LACASE PRECURSOR (EC	2.30e+02
421	5	17.2	448	4	015468	LEUCOCYTE INHIBITORY RE	2.30e+02	494	5	17.2	528	1	029668	CONSERVED HYPOTHETICAL	2.30e+02
422	5	17.2	449	5	P91028	COSMID C10G11.	2.30e+02	495	5	17.2	531	5	062026	CO4H5.3 PROTEIN.	2.30e+02
423	5	17.2	450	4	015636	TRANSCRIPTION FACTOR (	2.30e+02	496	5	17.2	532	8	Q21282	TRANSLATION INITIATION	2.30e+02
424	5	17.2	452	3	013636	HYPOTHETICAL 52.6 KD P	2.30e+02	497	5	17.2	533	3	060199	LACASE PRECURSOR.	2.30e+02
425	5	17.2	453	2	051881	ORF453 HYPOTHETICAL PR	2.30e+02	498	5	17.2	533	3	Q13490	CYTOCHROME P450.	2.30e+02
426	5	17.2	453	13	P87373	BMP5.	2.30e+02	499	5	17.2	534	10	096303	PHOSPHATE TRANSPORTER.	2.30e+02
427	5	17.2	455	10	043168	1-AMINOCYCLOPROPANE-1-	2.30e+02	500	5	17.2	536	2	084053	HYPOTHETICAL 56.8 KD P	2.30e+02
428	5	17.2	455	4	012875	ACTIVATED MET ONCOGENE	2.30e+02	501	5	17.2	538	3	07950	CHROMOSOME XII READING	2.30e+02
429	5	17.2	456	2	069454	P50 ADHESIN PRECURSOR	2.30e+02	502	5	17.2	544	2	083932	CHROMOSOME XII READING	2.30e+02
430	5	17.2	459	5	045080	C17H12.9 PROTEIN.	2.30e+02	503	5	17.2	545	10	P33690	HYPOTHETICAL 60.1 KD P	2.30e+02
431	5	17.2	460	2	Q34237	GUANYLYL TRANSFERASE (	2.30e+02	504	5	17.2	548	3	Q74447	SECY.	2.30e+02
432	5	17.2	460	5	093138	ORF1.	2.30e+02	505	5	17.2	548	2	Q32923	HYPOTHETICAL 63.0 KD P	2.30e+02
433	5	17.2	465	2	056267	GLUTAMATE SYNTHASE SMA	2.30e+02	506	5	17.2	551	10	Q38885	TPP-REQUIRING ENZYME.	2.30e+02
434	5	17.2	465	2	056621	GDP-MANNOSE PYROPHOSPH	2.30e+02	507	5	17.2	552	1	028554	SECY HOMOLOG.	2.30e+02
435	5	17.2	465	2	087163	ORF22-24 PROTEIN.	2.30e+02	508	5	17.2	553	8	063066	ACETOLACTATE SYNTHASE,	2.30e+02
436	5	17.2	466	10	082719	ACC SYNTHASE (EC 4.4.1	2.30e+02	509	5	17.2	554	5	076233	CPSECT.	2.30e+02
437	5	17.2	469	5	046123	GBR-2A PROTEIN PRECURS	2.30e+02	510	5	17.2	558	2	P96420	PUTATIVE SMALL G-PROTE	2.30e+02
438	5	17.2	470	5	069686	HYPOTHETICAL 53.0 KD P	2.30e+02	511	5	17.2	558	3	Q43119	PUTATIVE ALPHA-ISOPROP	2.30e+02
439	5	17.2	470	2	P95200	HYPOTHETICAL 50.4 KD P	2.30e+02	512	5	17.2	559	5	Q21416	PEPTIDE-N4-(N-ACETYL-B	2.30e+02
440	5	17.2	470	14	098199	MC031L.	2.30e+02	513	5	17.2	560	2	P71997	K10D3.2 PROTEIN.	2.30e+02
441	5	17.2	471	11	070419	CORTACTIN ISOFORM C.	2.30e+02	514	5	17.2	561	3	Q00062	HYPOTHETICAL 59.4 KD P	2.30e+02
442	5	17.2	471	10	Q39470	PUTATIVE CHROMOPLASTIC	2.30e+02	515	5	17.2	562	2	P76923	PUTATIVE PERMEASE.	2.30e+02
443	5	17.2	475	8	Q31802	RIBULOSE-1,5-BISPHOSPH	2.30e+02	516	5	17.2	565	2	069071	SIMILAR TO.	2.30e+02
444	5	17.2	476	5	023354	ZC504.1 PROTEIN.	2.30e+02	517	5	17.2	566	6	Q28970	OXIDOREDUCTASE CFA8.	2.30e+02
445	5	17.2	477	10	P94005	1-AMINOCYCLOPROPANE 1-	2.30e+02	518	5	17.2	569	3	Q14319	MYOSIN-VIIA MOTOR DOMA	2.30e+02
446	5	17.2	477	2	P95457	GLUTAMATE SYNTHASE SMA	2.30e+02	519	5	17.2	575	3	Q12198	HYPOTHETICAL 63.7 KD P	2.30e+02
447	5	17.2	477	5	045004	W03D8.1 PROTEIN.	2.30e+02	520	5	17.2	578	3	P78738	CHROMOSOME XII READING	2.30e+02
448	5	17.2	477	4	Q92731	ESTROGEN RECEPTOR BETA	2.30e+02	521	5	17.2	578	2	Q34810	EXCHITINASE.	2.30e+02
449	5	17.2	478	5	P92023	T10G3.7 PROTEIN.	2.30e+02	522	5	17.2	579	14	Q70153	HYPOTHETICAL 68.5 KD P	2.30e+02
450	5	17.2	478	5	017548	GLUCALPHA2B PROTEIN P	2.30e+02	523	5	17.2	581	5	Q22977	ENVELOPE GLYCOPROTEIN	2.30e+02
451	5	17.2	482	3	042656	HYPOTHETICAL PROTEIN C	2.30e+02	524	5	17.2	584	4	060366	COSMID F45F2.	2.30e+02
452	5	17.2	485	10	065475	HYPOTHETICAL 54.2 KD P	2.30e+02	525	5	17.2	585	3	P78978	HEPATOCTYTE GROWTH FACT	2.30e+02
453	5	17.2	485	5	018544	C39E9.7 PROTEIN.	2.30e+02	526	5	17.2	587	13	Q91250	REGULATOR OF XPR2 RIM1	2.30e+02
454	5	17.2	486	10	Q22724	F11F17.6 PROTEIN.	2.30e+02	527	5	17.2	590	10	065281	ESTROGEN RECEPTOR.	2.30e+02
455	5	17.2	489	5	Q93704	F42D1.3 PROTEIN.	2.30e+02	528	5	17.2	592	11	Q35567	ARABIDOPSIS THALIANA H	2.30e+02
456	5	17.2	491	10	Q43747	1-AMINOCYCLOPROPANE-1-	2.30e+02	529	5	17.2	592	10	081103	5-AMINOIMIDAZOLE-4-CAR	2.30e+02
457	5	17.2	492	5	002258	F38A1.7 PROTEIN.	2.30e+02	530	5	17.2	597	10	Q26060	POLYPHENOL OXIDASE PRE	2.30e+02
458	5	17.2	492	2	007488	SULFATE PERMEASE.	2.30e+02	531	5	17.2	600	5	Q21036	GLUCOSAMINE FRUCTOSE-6	2.30e+02
														CODED FOR BY C. ELEGAN	2.30e+02



532	5	17.2	500	2	085572	2.30e+02	PUTATIVE OLIGOPEPTIDE-
533	5	17.2	501	2	P71509	2.30e+02	METHANOL DEHYDROGENASE
534	5	17.2	503	2	P70041	2.30e+02	PENICILLIN-BINDING PRO
535	5	17.2	504	5	002374	2.30e+02	BRUNO.
536	5	17.2	506	3	P74764	2.30e+02	PUTATIVE ATP DEPENDENT
537	5	17.2	508	5	044991	2.30e+02	KO9H9.6 PROTEIN.
538	5	17.2	511	2	087321	2.30e+02	POLY(3-HYDROXYBUTYRATE
539	5	17.2	515	2	086781	2.30e+02	GLUCOSAMINE--FRUCTOSE-
540	5	17.2	520	14	089893	2.30e+02	(HHV-6).
541	5	17.2	520	3	Q12623	2.30e+02	GLUCOAMYLASE (EC 3.2.1
542	5	17.2	523	5	022976	2.30e+02	COSMID F45F2.
543	5	17.2	523	4	043841	2.30e+02	PAPS SYNTHASE.
544	5	17.2	524	4	043252	2.30e+02	PAPS SYNTHETASE.
545	5	17.2	524	4	P73332	2.30e+02	BIFUNCTIONAL ATP SULFU
546	5	17.2	524	11	Q05967	2.30e+02	ATP SULFURYLYASE-ADENOS
547	5	17.2	524	11	054820	2.30e+02	3'-PHOSPHOADENOSINE 5'
548	5	17.2	526	2	045877	2.30e+02	HA70.
549	5	17.2	531	3	014353	2.30e+02	PUTATIVE BIOTIN--PROTE
550	5	17.2	531	9	Q37967	2.30e+02	TELOMERASE.
551	5	17.2	533	10	081040	2.30e+02	F19D11.13 PROTEIN.
552	5	17.2	534	4	075114	2.30e+02	KIAA0616 PROTEIN (FRAG
553	5	17.2	537	14	070814	2.30e+02	POLYPROTEIN (FRAGMENT)
554	5	17.2	537	10	082612	2.30e+02	T5H22.5 PROTEIN.
555	5	17.2	540	10	024550	2.30e+02	MALATE DEHYDROGENASE (
556	5	17.2	540	3	053863	2.30e+02	PHOSPHOLIPASE B.
557	5	17.2	544	2	063677	2.30e+02	ALPHA-ISOPROPYL MALATE
558	5	17.2	551	11	Q03719	2.30e+02	POTASSIUM VOLTAGE GATE
559	5	17.2	552	3	Q93109	2.30e+02	REPELLENT PROTEIN.
560	5	17.2	557	5	017547	2.30e+02	GLUCALPHA2A PROTEIN.
561	5	17.2	560	3	013677	2.30e+02	HYPOTHETICAL 75.0 KD P
562	5	17.2	560	2	068460	2.30e+02	H+ TRANSLOCATING PYRO
563	5	17.2	561	1	052001	2.30e+02	SIMILAR TO EC0571.
564	5	17.2	565	1	015940	2.30e+02	UNC-14.
565	5	17.2	565	14	066473	2.30e+02	POLYPROTEIN (FRAGMENT)
566	5	17.2	566	5	019486	2.30e+02	F15H10.7 PROTEIN.
567	5	17.2	567	2	P71538	2.30e+02	PROBABLE CARBAMOYL-PHO
568	5	17.2	569	2	Q53445	2.30e+02	ENDOGLUCANASE 3 PRECUR
569	5	17.2	571	6	077676	2.30e+02	CGMP-DEPENDENT PROTEIN
570	5	17.2	571	4	Q13976	2.30e+02	CGMP-DEPENDENT PROTEIN
571	5	17.2	572	4	000319	2.30e+02	PAC CLONE DJ525N14.
572	5	17.2	573	4	060299	2.30e+02	KIAA0552 PROTEIN.
573	5	17.2	585	4	015271	2.30e+02	SYNAPTROPDIN.
574	5	17.2	592	5	076210	2.30e+02	HEXAMERIN A (FRAGMENT)
575	5	17.2	592	5	077461	2.30e+02	HEXAMERIN A (FRAGMENT)
576	5	17.2	593	5	077252	2.30e+02	GAMMA-GLUTAMYL-CYSTEINE
577	5	17.2	598	3	004563	2.30e+02	HYPOTHETICAL 80.5 KD P
578	5	17.2	598	4	015996	2.30e+02	LUTEINIZING HORMONE RE
579	5	17.2	599	11	Q63861	2.30e+02	SMOOTH MUSCLE MYOSIN H
580	5	17.2	701	4	Q14751	2.30e+02	LUTEINISING HORMONE-CH
581	5	17.2	702	2	Q49499	2.30e+02	HAEMAGGLUTININ HOMOLO
582	5	17.2	706	11	Q63862	2.30e+02	SMOOTH MUSCLE MYOSIN H
583	5	17.2	708	4	043641	2.30e+02	CACO-2 OLIGOPEPTIDE TR
584	5	17.2	718	10	Q38887	2.30e+02	OVULE-SPECIFIC HOMEORO
585	5	17.2	719	2	Q51912	2.30e+02	PROTEIN L PRECURSOR.
586	5	17.2	720	14	Q02568	2.30e+02	GENOME POLYPROTEIN (CO
587	5	17.2	720	2	066880	2.30e+02	CADPSID PROTEIN (NAD DEPEND
588	5	17.2	720	14	084771	2.30e+02	CADPSID PROTEIN PRECUR
589	5	17.2	730	10	022989	2.30e+02	CELLULOSE SYNTHASE ISO
590	5	17.2	731	14	Q00279	2.30e+02	GENOME POLYPROTEIN
591	5	17.2	731	14	Q65090	2.30e+02	P1 POLYPROTEIN (FRAGME
592	5	17.2	731	2	070039	2.30e+02	PENICILLIN-BINDING PRO
593	5	17.2	731	11	088327	2.30e+02	ALPHA-CATENIN RELATED
594	5	17.2	732	14	Q00283	2.30e+02	GENOME POLYPROTEIN
595	5	17.2	732	14	Q65066	2.30e+02	CADPSID PROTEIN P1 (FRA
596	5	17.2	732	14	Q65089	2.30e+02	P1 POLYPEPTIDE (FRAGME
597	5	17.2	732	14	Q01119	2.30e+02	GENOME POLYPROTEIN
598	5	17.2	733	14	Q01120	2.30e+02	GENOME POLYPROTEIN
599	5	17.2	734	4	076084	2.30e+02	ALPHA-CATENIN-LIKE PRO
600	5	17.2	736	14	P87677	2.30e+02	POLYPROTEIN (FRAGMENT)
601	5	17.2	741	11	Q61285	2.30e+02	PEROXISOMAL PROTEIN AL
602	5	17.2	745	10	P23041	2.30e+02	HOMODOMAIN PROTEIN AH
603	5	17.2	746	10	Q23067	2.30e+02	SIMILARITY TO HAEMOPHI
604	5	17.2	746	5	Q23066	2.30e+02	SIMILAR TO THE CARNITI
605	5	17.2	749	10	033045	2.30e+02	SIMILARITY TO PHALAENO
606	5	17.2	751	2	007850	2.30e+02	VAHL.
607	5	17.2	758	14	Q65106	2.30e+02	FIGURE 2.
608	5	17.2	759	14	Q00574	2.30e+02	GENOME POLYPROTEIN (CO
609	5	17.2	766	10	023538	2.30e+02	RESISTANCE GENE HOMOLO
610	5	17.2	766	10	040988	2.30e+02	HOMEOBOX PROTEIN.
611	5	17.2	771	11	070570	2.30e+02	POLYMERIC-IMMUNOGLOBUL
612	5	17.2	774	5	045134	2.30e+02	C05D2.5 PROTEIN.
613	5	17.2	774	5	Q17950	2.30e+02	C14A4.2 PROTEIN.
614	5	17.2	778	14	Q84509	2.30e+02	PRO-RICH.
615	5	17.2	782	14	Q98704	2.30e+02	NIB PROTEIN (FRAGMENT)
616	5	17.2	783	4	075176	2.30e+02	KIAA0692 PROTEIN (FRAG
617	5	17.2	788	2	P72224	2.30e+02	QUINOLINE 2-OXIDOREduc
618	5	17.2	788	5	018510	2.30e+02	INSECT INTESTINAL MUCI
619	5	17.2	791	4	051220	2.30e+02	KIAA0625 PROTEIN (FRAG
620	5	17.2	792	5	061840	2.30e+02	F5066.1 PROTEIN (FRAGM
621	5	17.2	796	2	025163	2.30e+02	BIOTIN SULFOXIDE REDUC
622	5	17.2	798	2	050255	2.30e+02	STONATIN LIKE PROTEIN
623	5	17.2	801	14	Q85437	2.30e+02	VIRAL STRUCTURAL PROTE
624	5	17.2	806	2	Q55838	2.30e+02	SENSORY TRANSDUCTION H
625	5	17.2	806	6	018783	2.30e+02	PLASMINOGEN.
626	5	17.2	807	5	018511	2.30e+02	INSECT INTESTINAL MUCI
627	5	17.2	807	5	077341	2.30e+02	MAL3P4.5 PROTEIN.
628	5	17.2	808	5	018409	2.30e+02	TESTIS-SPECIFIC RNP-TY
629	5	17.2	809	5	P90534	2.30e+02	RSC12 (FRAGMENT).
630	5	17.2	817	5	Q94232	2.30e+02	SIMILARITY TO NEUTRAL
631	5	17.2	818	14	056051	2.30e+02	STRUCTURAL POLYPROTEIN
632	5	17.2	822	11	054923	2.30e+02	RSEC15.
633	5	17.2	822	2	066172	2.30e+02	TRBE.
634	5	17.2	822	2	087117	2.30e+02	DNA GYRASE A SUBUNIT.
635	5	17.2	824	4	Q16349	2.30e+02	H-NUC.
636	5	17.2	825	5	Q23484	2.30e+02	COSMID ZK418.
637	5	17.2	828	4	Q92738	2.30e+02	PROTEIN RELATED N-TERN
638	5	17.2	835	2	069873	2.30e+02	HYPOTHETICAL 89.0 KD P
639	5	17.2	835	2	056012	2.30e+02	CELL-SURFACE-ASSOCIATE
640	5	17.2	840	5	061894	2.30e+02	T07H8.6 PROTEIN.
641	5	17.2	841	14	Q67632	2.30e+02	UL9 HOMOLOG.
642	5	17.2	843	14	Q74109	2.30e+02	DNA POLYMERASE (EC 2.7
643	5	17.2	845	14	Q81096	2.30e+02	ENVELOPE GLYCOPROTEIN
644	5	17.2	848	14	040366	2.30e+02	ENVELOPE GLYCOPROTEIN
645	5	17.2	848	5	Q94419	2.30e+02	ZK858.1 PROTEIN.
646	5	17.2	852	14	Q74112	2.30e+02	ENVELOPE GLYCOPROTEIN
647	5	17.2	852	14	Q41546	2.30e+02	ENVELOPE GLYCOPROTEIN
648	5	17.2	852	14	Q89607	2.30e+02	PUTATIVE PHEROMONE REC
649	5	17.2	855	11	070409	2.30e+02	ENVELOPE GLYCOPROTEIN
650	5	17.2	856	14	040364	2.30e+02	ENVELOPE GLYCOPROTEIN
651	5	17.2	857	14	040357	2.30e+02	ENVELOPE GLYCOPROTEIN
652	5	17.2	858	6	018868	2.30e+02	DELAYED RECTIFIER POT
653	5	17.2	858	14	Q74114	2.30e+02	GP160.
654	5	17.2	860	14	Q93091	2.30e+02	ENVELOPE GLYCOPROTEIN
655	5	17.2	860	14	Q74110	2.30e+02	GP160.
656	5	17.2	860	14	Q74111	2.30e+02	ENVELOPE GLYCOPROTEIN
657	5	17.2	861	14	Q90599	2.30e+02	ENVELOPE GLYCOPROTEIN
658	5	17.2	862	14	Q76606	2.30e+02	ENVELOPE GLYCOPROTEIN
659	5	17.2	862	14	Q74113	2.30e+02	GP160.
660	5	17.2	863	14	Q40360	2.30e+02	ENVELOPE GLYCOPROTEIN
661	5	17.2	863	14	077377	2.30e+02	ENVELOPE PROTEIN GP120
662	5	17.2	865	14	P88147	2.30e+02	ENVELOPE POLYPROTEIN.
663	5	17.2	866	14	Q76868	2.30e+02	ENVELOPE POLYPROTEIN.
664	5	17.2	866	14	Q74108	2.30e+02	GP160.
665	5	17.2	869	14	Q76627	2.30e+02	ENVELOPE POLYPROTEIN.
666	5	17.2	871	14	Q57074	2.30e+02	ENVELOPE GLYCOPROTEIN.
667	5	17.2	872	11	055073	2.30e+02	ENVELOPE GLYCOPROTEIN.
668	5	17.2	872	11	055000	2.30e+02	HYPOTHETICAL 92.8 KD P
669	5	17.2	874	14	Q40460	2.30e+02	ENV PROTEIN (FRAGMENT)
670	5	17.2	875	10	Q48930	2.30e+02	NITRATE REDUCTASE (FRA
671	5	17.2	876	14	Q79670	2.30e+02	ENVELOPE PROTEIN GP120
672	5	17.2	877	14	Q76576	2.30e+02	ENVELOPE GENE (FRAGMEN
673	5	17.2	878	14	Q83101	2.30e+02	RNA-DEPENDENT RNA POLY
674	5	17.2	881	5	015307	2.30e+02	T21H3.2 PROTEIN.
675	5	17.2	881	10	Q23866	2.30e+02	ME12-LIKE PROTEIN.
676	5	17.2	891	5	Q93763	2.30e+02	F53C11.5 PROTEIN.
677	5	17.2	891	14	Q66018	2.30e+02	GB GLYCOPROTEIN.



824	5	17.2	4993	14	011993	CYTOPATHIC GENOMIC RNA	2.30e+02	897	4	13.8	395	2	Q55456	HYPOTHETICAL 43.4 KD P	5.59e+03
825	5	17.2	26926	4	Q10466	TITIN, HEART ISOFORM N	2.30e+02	898	4	13.8	398	5	O62169	F14F4.1 PROTEIN.	5.59e+03
826	4	13.8	66	5	Q25652	HISTIDINE-RICH KNOB PR	5.59e+03	899	4	13.8	407	5	P91369	COSMID K1H12.	5.59e+03
827	4	13.8	96	1	Q73991	96AA LONG HYPOTHETICAL	5.59e+03	900	4	13.8	408	2	Q74265	HYPOTHETICAL 45.3 KD P	5.59e+03
828	4	13.8	98	2	Q54318	PERIPLASMIC TRANSPORT	5.59e+03	901	4	13.8	418	2	O84014	HYPOTHETICAL 48.1 KD P	5.59e+03
829	4	13.8	103	5	Q02619	ZK131.4 PROTEIN.	5.59e+03	902	4	13.8	419	5	Q77287	SU(P) PROTEIN.	5.59e+03
830	4	13.8	106	10	Q41633	HYPOTHETICAL 11.8 KD P	5.59e+03	903	4	13.8	424	4	Q15224	PREGNANCY-SPECIFIC-BET	5.59e+03
831	4	13.8	118	8	Q96005	NADH DEHYDROGENASE SUB	5.59e+03	904	4	13.8	428	5	O16134	TRANSFORMING GROWTH FA	5.59e+03
832	4	13.8	118	8	Q96030	NADH DEHYDROGENASE SUB	5.59e+03	905	4	13.8	429	2	Q52000	DICARBOXYLIC ACID TRAN	5.59e+03
833	4	13.8	127	2	Q26075	CONSERVED HYPOTHETICAL	5.59e+03	906	4	13.8	447	5	Q21147	SIMILAR TO EPOXYHETICA	5.59e+03
834	4	13.8	131	7	Q46869	MHC CLASS II BETA CHAI	5.59e+03	907	4	13.8	449	1	O58175	449AA LONG HYPOTHETICA	5.59e+03
835	4	13.8	133	1	Q59202	133AA LONG HYPOTHETICA	5.59e+03	908	4	13.8	461	1	O29450	GLYCOLATE OXIDASE SUBU	5.59e+03
836	4	13.8	137	8	Q36996	ORF137.	5.59e+03	909	4	13.8	464	2	P73249	HYPOTHETICAL 51.6 KD P	5.59e+03
837	4	13.8	141	2	Q44784	ORF-10.	5.59e+03	910	4	13.8	469	2	P75040	MG061 HOMOLOG.	5.59e+03
838	4	13.8	141	2	Q44779	ORF-10.	5.59e+03	911	4	13.8	469	2	P73738	HYPOTHETICAL 52.5 KD P	5.59e+03
839	4	13.8	142	2	P73568	HYPOTHETICAL 15.8 KD P	5.59e+03	912	4	13.8	471	5	O62133	F02H6.5 PROTEIN.	5.59e+03
840	4	13.8	154	1	O58327	154AA LONG HYPOTHETICA	5.59e+03	913	4	13.8	477	3	O42968	HYPOTHETICAL 54.9 KD P	5.59e+03
841	4	13.8	163	3	Q00060	PEPTIDYL-PROLYL CIS/TR	5.59e+03	914	4	13.8	479	2	O86001	SALICYLALDEHYDE DHYDR	5.59e+03
842	4	13.8	173	5	Q25666	INTEGRAL MEMBRANE PROT	5.59e+03	915	4	13.8	480	4	O75472	TUMOROUS IMAGINAL DISC	5.59e+03
843	4	13.8	174	2	O86042	ATP/GTP BINDING PROTEI	5.59e+03	916	4	13.8	485	2	P71815	HYPOTHETICAL 52.0 KD P	5.59e+03
844	4	13.8	178	2	O52707	ATTENUATION REGULATORY	5.59e+03	917	4	13.8	490	4	O14670	TEB4 PROTEIN.	5.59e+03
845	4	13.8	180	4	O77066	SGG10-LIKE-PROTEIN.	5.59e+03	918	4	13.8	495	2	O85370	PREPROAEROLYSIN.	5.59e+03
846	4	13.8	182	2	O85527	Z61F PROTEIN (FRAGMENT	5.59e+03	919	4	13.8	495	5	O94224	SIMILARITY TO IG-LIKE	5.59e+03
847	4	13.8	187	2	O87021	Z2R PROTEIN (FRAGMENT)	5.59e+03	920	4	13.8	497	2	O67661	TRANSCRIPTIONAL REGULA	5.59e+03
848	4	13.8	192	2	O34384	YCEE.	5.59e+03	921	4	13.8	500	2	O52526	POLYSACCHARIDE EXPORT	5.59e+03
849	4	13.8	221	10	O64464	BETA 6 SUBUNIT OF 20S	5.59e+03	922	4	13.8	503	4	P78324	SHP SUBSTRATE-1 PRECUR	5.59e+03
850	4	13.8	222	2	O08368	GLUTATHIONE PEROXIDASE	5.59e+03	923	4	13.8	509	2	O31149	ENZYME I (FRAGMENT).	5.59e+03
851	4	13.8	227	1	Q05745	ORF AND MOAA GENE.	5.59e+03	924	4	13.8	517	2	P73549	HYPOTHETICAL 53.8 KD P	5.59e+03
852	4	13.8	232	8	Q34238	POT. ALT. PETB GENE PR	5.59e+03	925	4	13.8	522	2	P73982	HYPOTHETICAL 61.7 KD P	5.59e+03
853	4	13.8	235	2	O53324	OXIDOREDUCTASE.	5.59e+03	926	4	13.8	522	4	O99592	TRANSILIN ASSOCIATED ZI	5.59e+03
854	4	13.8	240	2	O33022	RIBONUCLEASE HII.	5.59e+03	927	4	13.8	531	8	O47582	NADH DEHYDROGENASE SUB	5.59e+03
855	4	13.8	244	10	O65156	GLUTATHIONE PEROXIDASE	5.59e+03	928	4	13.8	533	2	O52729	BETA-XILOSIDASE (EC 3.	5.59e+03
856	4	13.8	244	2	O34170	FLGF.	5.59e+03	929	4	13.8	538	4	O92692	POLIO VIRUS RECEPTOR P	5.59e+03
857	4	13.8	247	5	P91202	COSMID ECAP9.	5.59e+03	930	4	13.8	539	2	O66190	60 KD CHAPERONIN (PROT	5.59e+03
858	4	13.8	250	2	O87659	PERIPLASMIC FIMBRIAL C	5.59e+03	931	4	13.8	540	2	O66194	60 KD CHAPERONIN (PROT	5.59e+03
859	4	13.8	256	1	O58253	256AA LONG HYPOTHETICA	5.59e+03	932	4	13.8	542	10	O64858	F411.4 PROTEIN.	5.59e+03
860	4	13.8	256	2	O33539	POTATIVE REGULATOR OF	5.59e+03	933	4	13.8	544	2	O50305	GROEL.	5.59e+03
861	4	13.8	256	2	P70816	OUTER SURFACE PROTEIN	5.59e+03	934	4	13.8	553	10	O04559	SIMILAR TO HUMAN LINE-	5.59e+03
862	4	13.8	261	8	Q36702	ATP SYNTHASE A CHAIN (	5.59e+03	935	4	13.8	553	10	O04559	T7N9.12.	5.59e+03
863	4	13.8	268	4	P78346	RNASEP PROTEIN P30.	5.59e+03	936	4	13.8	557	5	O21837	COSMID R08C7.	5.59e+03
864	4	13.8	268	2	O25169	TRANSPORASE-LIKE PROTE	5.59e+03	937	4	13.8	569	5	O25970	MAJOR MEROZOITE SURFAC	5.59e+03
865	4	13.8	268	2	O53559	PURINE NUCLEOSIDE PHOS	5.59e+03	938	4	13.8	570	5	O25968	MAJOR MEROZOITE SURFAC	5.59e+03
866	4	13.8	273	2	O54324	OUTER SURFACE PROTEIN	5.59e+03	939	4	13.8	574	4	O75134	KIAA0639 PROTEIN (FRAG	5.59e+03
867	4	13.8	276	5	O18443	CHYMOTRYPSIN-LIKE PROT	5.59e+03	940	4	13.8	576	2	P71664	HYPOTHETICAL GLYCINE-R	5.59e+03
868	4	13.8	278	2	O59823	RNA POLYMERASE SIGMA F	5.59e+03	941	4	13.8	577	10	O48768	T21L14.9 PROTEIN.	5.59e+03
869	4	13.8	285	2	P73514	HYPOTHETICAL 32.0 KD P	5.59e+03	942	4	13.8	581	6	O45651	PROLACTIN RECEPTOR LON	5.59e+03
870	4	13.8	281	10	O22239	SIMILAR TO GLYCYL TRNA	5.59e+03	943	4	13.8	581	5	O02365	CYTOSOLASMIC INTERMEDIA	5.59e+03
871	4	13.8	287	5	P91997	F53F1.2 PROTEIN.	5.59e+03	944	4	13.8	591	8	O36149	NADH-UBIQUINONE OXIDOR	5.59e+03
872	4	13.8	297	5	O16786	T21D12.3 PROTEIN.	5.59e+03	945	4	13.8	595	2	O86616	BETA-GALACTOSIDASE.	5.59e+03
873	4	13.8	297	2	O51629	REP PROTEIN.	5.59e+03	946	4	13.8	607	2	O51476	DNA POLYMERASE III TAU	5.59e+03
874	4	13.8	300	2	O34574	YEFB PROTEIN.	5.59e+03	947	4	13.8	608	2	O45998	2C15.7 PROTEIN.	5.59e+03
875	4	13.8	301	2	O07156	HYPOTHETICAL 32.9 KD P	5.59e+03	948	4	13.8	633	5	O18270	ERYTHROID MEMBRANE PRO	5.59e+03
876	4	13.8	311	2	O51632	REP PROTEIN.	5.59e+03	949	4	13.8	641	4	O14245	PROTEIN S (FRAGMENT).	5.59e+03
877	4	13.8	313	5	O20166	F38E11.3 PROTEIN.	5.59e+03	950	4	13.8	648	6	O29094	BETA-GLUCURONIDASE (EC	5.59e+03
878	4	13.8	316	4	O60218	ALDOSE REDUCTASE-LIKE	5.59e+03	951	4	13.8	648	6	O77695	BETA-GLUCURONIDASE (EC	5.59e+03
879	4	13.8	330	5	Q25025	CYSTEINE PROTEINASE.	5.59e+03	952	4	13.8	653	4	O14829	PROTEIN PHOSPHATASE WI	5.59e+03
880	4	13.8	336	4	O14100	HETEROGENEOUS NUCLEAR	5.59e+03	953	4	13.8	667	2	Q55462	NITRATE TRANSPORT PROT	5.59e+03
881	4	13.8	343	2	O54304	LYSINE CYCLODEAMINASE.	5.59e+03	954	4	13.8	670	4	O92873	B-CELL RECEPTOR CD22-A	5.59e+03
882	4	13.8	344	2	O50325	ENDO-PECTATE LYASE PRE	5.59e+03	955	4	13.8	679	1	O29619	CONSERVED HYPOTHETICAL	5.59e+03
883	4	13.8	348	2	P95344	OUTER MEMBRANE PROTEIN	5.59e+03	956	4	13.8	683	2	O52780	XYLANASE U.	5.59e+03
884	4	13.8	348	2	P95347	OUTER MEMBRANE PROTEIN	5.59e+03	957	4	13.8	702	5	O46065	COSMID 17A9.	5.59e+03
885	4	13.8	350	6	O19069	SUCCINYL-COA SYNTHETAS	5.59e+03	958	4	13.8	707	2	O05415	TNPX.	5.59e+03
886	4	13.8	350	5	O61093	HYPOTHETICAL 39.4 KD P	5.59e+03	959	4	13.8	707	6	P72207	BETA-SPECTRIN (FRAGMEN	5.59e+03
887	4	13.8	357	2	P75470	F10. ORF357 PROTEIN.	5.59e+03	960	4	13.8	716	2	P72201	ORF1, ORF2 AND ORF3 GE	5.59e+03
888	4	13.8	359	2	O50440	HYPOTHETICAL 37.8 KD P	5.59e+03	961	4	13.8	798	5	O17484	SIMILARITY TO D. MELAN	5.59e+03
889	4	13.8	359	5	Q26596	SER- AND THR-RICH PROT	5.59e+03	962	4	13.8	836	2	P76450	FROM BASES 2276392 TO	5.59e+03
890	4	13.8	361	2	P96391	HYPOTHETICAL 38.1 KD P	5.59e+03	963	4	13.8	837	5	O17968	CI4BL1.6 PROTEIN.	5.59e+03
891	4	13.8	365	8	O02687	HYPOTHETICAL 42.0 KD P	5.59e+03	964	4	13.8	884	2	O83703	PENICILLIN-BINDING PRO	5.59e+03
892	4	13.8	365	2	O53365	ADENOSINE DEAMINASE.	5.59e+03	965	4	13.8	884	4	O14953	PROTEIN-TYROSINE KINAS	5.59e+03
893	4	13.8	366	2	O53033	TRANSPORASE.	5.59e+03	966	4	13.8	942	5	O18298	C29E6.4 PROTEIN.	5.59e+03
894	4	13.8	367	2	O48524	MOBILIZATION PROTEIN.	5.59e+03	967	4	13.8	946	5	Q23399	SIMILAR TO RAT GAP-ASS	5.59e+03
895	4	13.8	377	2	O83323	CONSERVED HYPOTHETICAL	5.59e+03	968	4	13.8	947	2	O86487	SDRC PROTEIN.	5.59e+03
896	4	13.8	380	3	Q06791	ORF (FRAGMENT).	5.59e+03	969	4	13.8	956	5	Q18129	SIMILAR TO DROSOPHILA	5.59e+03

970 4 13.8 958 4 Q13592 B219/OB RECEPTOR ISOFO 5.59e+03  
 971 4 13.8 964 4 Q14690 INTEGRIN VARIANT BETA4 5.59e+03  
 972 4 13.8 974 6 Q62852 CAPACITATIVE CALCIUM E 5.59e+03  
 973 4 13.8 1000 2 Q85082 LACTOFERRIN BINDING PR 5.59e+03  
 974 4 13.8 1036 2 Q86999 S-LAYER PROTEIN. 5.59e+03  
 975 4 13.8 1070 2 Q44070 NUCLEASE. 5.59e+03  
 976 4 13.8 1070 5 Q94290 CODED FOR BY C. ELEGAN 5.59e+03  
 977 4 13.8 1081 2 Q68831 SURFACE ANTIGEN BSPA. 5.59e+03  
 978 4 13.8 1139 2 Q54073 ANCHOR PROTEIN, LCM. 5.59e+03  
 979 4 13.8 1179 2 Q03744 139 KD CRYSTAL PROTEIN 5.59e+03  
 980 4 13.8 1198 4 Q75882 ATTRACTIN. 5.59e+03  
 981 4 13.8 1306 2 Q47766 AGGREGATION SUBSTANCE 5.59e+03  
 982 4 13.8 1385 2 Q45760 DELTA-ENDOTOXIN (FRAGM 5.59e+03  
 983 4 13.8 1389 5 Q77319 MALP3.1 PROTEIN. 5.59e+03  
 984 4 13.8 1418 6 Q28396 TYPE II COLLAGEN. 5.59e+03  
 985 4 13.8 1481 6 Q62673 CYSTIC FIBROSIS TRANSM 5.59e+03  
 986 4 13.8 1752 4 Q5341 BETA4- INTEGRIN. 5.59e+03  
 987 4 13.8 1755 3 P87195 CHROMOSOME X READING F 5.59e+03  
 988 4 13.8 1805 4 Q92545 MYELOBLAST KIAA0257 (F 5.59e+03  
 989 4 13.8 1867 3 Q13967 PUTATIVE 1,3-BETA-GLUC 5.59e+03  
 990 4 13.8 1953 3 Q13450 'PPF3P'. 5.59e+03  
 991 4 13.8 1978 2 P73551 HYPOTHETICAL 213.8 KD 5.59e+03  
 992 4 13.8 2019 5 Q8292 ZK1151.3 PROTEIN. 5.59e+03  
 993 4 13.8 2037 5 Q22511 COSMID T14G12. 5.59e+03  
 994 4 13.8 2277 5 Q23187 W06A7.3A PROTEIN (FRAG 5.59e+03  
 995 4 13.8 2543 2 Q31784 POLYKETIDE SYNTHASE. 5.59e+03  
 996 4 13.8 2894 1 Q58791 HYPOTHETICAL PROTEIN M 5.59e+03  
 997 4 13.8 3016 2 P73590 HYPOTHETICAL 311.5 KD 5.59e+03  
 998 4 13.8 3229 5 Q26912 PROTEIN 1 OF A DISPERS 5.59e+03  
 999 4 13.8 4574 4 Q16640 PLECTIN. 5.59e+03  
 1000 4 13.8 13288 6 Q18758 SUBMAXILLARY APOMUCIN. 5.59e+03

## ALIGNMENTS

RESULT 1  
 ID Q20433 PRELIMINARY; PRT; 67 AA.  
 AC Q20433;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE COSMID F45E12.  
 GN F45E12.6.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEIA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSOUGH R., ANDERSON K., RAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL NATURE 368:32-38(1994).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX LATREILLE P.;  
 RA LATREILLE P.;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX WATERSTON R.;  
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; U29536; G868265; -  
 SQ SEQUENCE 67 AA; 7865 MW; ECEA86AD CRC32;  
 Query Match 20.7%; Score 6; DB 5; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 21 FPSPLV 26  
 |||||  
 Qy 14 FPSPLV 19  
 |||||  
 RESULT 2  
 ID P75213 PRELIMINARY; PRT; 152 AA.  
 AC P75213;  
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE H03\_ORF152 PROTEIN.  
 OS MYCOPLASMA PNEUMONIAE.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;  
 OC MYCOPLASMATACEAE; MYCOPLASMA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-M129;  
 RX MEDLINE; 97105885.  
 RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.C.,  
 RA HERRMANN R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-M129;  
 RA HIMMELREICH R., HILBERT H., LI B.C.;  
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AE00027; G1673948; -  
 SQ SEQUENCE 152 AA; 17463 MW; 08EA6E52 CRC32;  
 Query Match 20.7%; Score 6; DB 2; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 20 PLVVA 25  
 |||||  
 Qy 17 PLVVA 22  
 |||||  
 RESULT 3  
 ID Q26372 PRELIMINARY; PRT; 205 AA.  
 AC Q26372;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE ACETYL / ACYL TRANSFERASE RELATED PROTEIN.  
 GN MTH272.  
 OS METHANOBACTERIUM THERMAUTOTROPHICUM.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;  
 OC METHANOBACTERIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE; 98037514.  
 RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
 RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
 RA HARRISON D., HOANG L., KEAGLE P., LUM W., POTIER B., QIU D.,  
 RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
 RA JEWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
 RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,  
 RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. BACTERIOL. 179:7135-7155(1997).

```

DR EMBL: AF000813; G2621323; -.
DR PFM: PF00132; hexapep; 4.
KW TRANSFERASE.
SQ SEQUENCE 205 AA; 22486 MW; 66F2FEC3 CRC32;

Query Match          20.7%; Score 6; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 146 IROGAS 151
QY 24 IROGAS 29

RESULT 4
ID Q44479 PRELIMINARY; PRT; 225 AA.
AC Q44479;
DT 01-JUN-1998 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE ORF225.
OS ANABAENA VARIABILIS.
OC BACTERIA; CYANOBACTERIA; NOSTOCALES; NOSTOCACEAE; ANABAENA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R3;
RA LOS D.A., MURATA N.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U41665; G1136328; -.
SQ SEQUENCE 225 AA; 25585 MW; 8FD9E166 CRC32;

Query Match          20.7%; Score 6; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 216 PLVVVA 221
QY 17 PLVVVA 22

RESULT 5
ID Q49493 PRELIMINARY; PRT; 249 AA.
AC Q49493;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 27.2 KD PROTEIN.
GN F28A23.120.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RA BEVAN M., WEICHSSELGARTNER M., FARTMANN B., GRANDERATH K., DAUNER D.,
RA HERZL A., NEUMANN S., HOEISEL J., JESSE T., HEIJNEN L., VOS P.,
RA MEWES H.W., MAYER K., SCHUELLER C.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AL021961; E1253338; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 249 AA; 27177 MW; CC5E0363 CRC32;

Query Match          20.7%; Score 6; DB 10; Length 249;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 203 PSLVV 208
QY 15 PSLVV 20

RESULT 6
ID Q70018 PRELIMINARY; PRT; 304 AA.

```

---

```

AC 070018;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 32.8 KD PROTEIN.
GN SHUT.
OS SHIGELLA DYSENTERIAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC SHIGELLA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O-4576;
RX MEDLINE; 95286477.
RA MILLS M., PAYNE S.M.;
RT "Genetics and regulation of heme iron transport in Shigella
RT dysenteriae and detection of an analogous system in Escherichia coli
RT O157:H7."
RL J. BACTERIOL. 177:3004-3009(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O-4576;
RX MEDLINE; 98053990.
RA MILLS M., PAYNE S.M.;
RT "Identification of shuA, the gene encoding the heme receptor of
RT Shigella dysenteriae, and analysis of invasion and intracellular
RT multiplication of a shuA mutant."
RL INFECT. IMMUN. 65:5358-5363(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-O-4576;
RA WYCKOFF E.E., DUNCAN D., TORRES A.G., MILLS M., MAASE K., PAYNE S.M.;
RL MOL. MICROBIOL. 0:0-0(1998).
DR EMBL: U64516; G2967538; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 304 AA; 32809 MW; 6DC0654C CRC32;

Query Match          20.7%; Score 6; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 147 YANIRQ 152
QY 21 YANIRQ 26

RESULT 7
ID P88951 PRELIMINARY; PRT; 305 AA.
AC P88951;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF 60.
OS KAPOSI'S SARCOMA-ASSOCIATED HERPESVIRUS.
OC VIRUSES; DSDNA VIRUSES; NO RNA STAGE; HERPESVIRIDAE;
OC GAMMAHERPESVIRINAE; RHADINOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97121480.
RA RUSSO J.J., BOHENZKY R.A., CHIEN M.C., CHEN J., YAN M., MADDALENA D.,
RA PARRY J.P., PERUZZI D., EDELMAN I.S., CHANG Y., MOORE P.S.;
RT "Nucleotide sequence of the kaposi sarcoma-associated herpesvirus
RT (HHV8)."
RL PROC. NATL. ACAD. SCI. U.S.A. 93:14862-14867(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97094384.
RA MOORE P.S., BASHOFF C., WEISS R.A., CHANG Y.;
RT "Molecular mimicry of human cytokine and cytokine response pathway
RT genes by KSHV."
RL SCIENCE 274:1739-1744(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97296220.

```

RA NEIPEL F., ALBRECHT J.C., FLECKENSTEIN B.;  
 RT "Cell-homologous genes in the Kaposi's sarcoma-associated  
 RT rhadinovirus human herpesvirus 8: determinants of its  
 RT pathogenicity?";  
 RL J. VIROL. 71:4187-4192(1997).  
 DR EMBL; U75698; G1718316; -;  
 DR EMBL; U93872; G2246534; -;  
 DR PFAM; PF00268; ribonuc\_red; 1.  
 SQ SEQUENCE 305 AA; 34998 MW; 22D8C7A1 CRC32;

Query Match 20.7%; Score 6; DB 14; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 VYVANI 104  
 Qy 19 VYVANI 24

RESULT 8 PRELIMINARY; PRT; 321 AA.  
 ID O32281  
 AC O32281  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE CYTOCHROME A3 QUINOL OXIDASE (SUBUNIT II).  
 GN OXAA.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE; 98044033.

RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
 RA GHM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUTSEPPI G., GUY B.J., HAGA K., HAIECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
 RA JORIS B., KARAIYA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NOONE D., O'REILLY M., OGAWA K., OGIMURA A., OUDEGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTELELE D., PORWOLLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SEKIGUCHI J., SEKONKA A., SERO S.J., SERROR P., SHIN B.S., SOLDI B.,  
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUT R., WEDLER E., WEDLER H., WEITZENEGER T.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL NATURE 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; 299123; E1186316; -;  
 SQ SEQUENCE 321 AA; 36253 MW; 38F57227 CRC32;

Query Match 20.7%; Score 6; DB 2; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 130 PLVYVA 135  
 Qy 17 PLVYVA 22

RESULT 9 PRELIMINARY; PRT; 342 AA.  
 ID O13386  
 AC O13386;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE PROTEASOME REGULATORY SUBUNIT 12.  
 GN PRS12.  
 OS HYPOCREA JECORINA.  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;  
 OC HYPOCREALES; HYPOCREACEAE; HYPOCREA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-QM 9414;  
 RA GOLLER S.P., GORFER M., KUBICEK C.P.;  
 RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF028783; G2599117; -;  
 KW PROTEASOME.  
 SQ SEQUENCE 342 AA; 38312 MW; AC8BDA24 CRC32;

Query Match 20.7%; Score 6; DB 3; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 TVTSKT 38  
 Qy 2 TVTSKT 7

RESULT 10 PRELIMINARY; PRT; 363 AA.  
 ID Q93193  
 AC Q93193;  
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)  
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE C07A4.3 PROTEIN.  
 GN C07A4.3.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNITEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA WHITE S.;  
 RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL NATURE 368:32-38(1994).  
 DR EMBL; Z78536; E1343817; -;  
 SQ SEQUENCE 363 AA; 41472 MW; 05A0DA44 CRC32;

Query Match 20.7%; Score 6; DB 5; Length 363;

```

Best Local Similarity 100.0%; Pred. No. 5.03e+00; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Db 262 SKPSP 267
    |||||
QY 12 SKPSP 17

RESULT 11
ID Q66726 PRELIMINARY; PRT: 365 AA.
AC Q66726;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GAG PROTEIN (FRAGMENT).
GN GAG.
OS EQUINE INFECTIOUS ANEMIA VIRUS (EIAV).
OC VIRUSES; RETROID VIRUSES; RETROVIRIDAE; LENTIVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92292230.
RA PERRY S.T., FLAHERTY M.T., KELLEY M.J., CLABOUGH D.L., TRONICK S.R.,
RA COGGINS L., WHETTER L., LENGEL C.R., FULLER F.;
RT "The surface envelope protein gene region of equine infectious anemia
RT virus is not an important determinant of tropism in vitro.";
RL J. VIROL. 66:4085-4097(1992).
DR EMBL; M87575; G554639; -.
DR PFAM; PF00507; gag_p24; 1.
KW POLYPROTEIN.
FT NON_TER 365
SQ SEQUENCE 365 AA; 41300 MW; A1899957 CRC32;

Query Match 20.7%; Score 6; DB 14; Length 365;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 276 NIRQGA 281
    |||||
QY 23 NIRQGA 28

RESULT 12
ID Q07110 PRELIMINARY; PRT: 385 AA.
AC Q07110;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CELL DIVISION PROTEIN.
GN DIVIB.
OS ENTEROCOCCUS FAECALIS (STREPTOCOCCUS FAECALIS).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; ENTEROCOCCACEAE;
OC ENTEROCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-A24836;
RX MEDLINE; 97431524.
RA PUCCI M.J., THANASSI J.A., DISCOTTO L.F., KESSLER R.E.,
RA DOUGHERTY T.J.;
RT "Identification and characterization of cell wall-cell division gene
RT clusters in pathogenic gram-positive cocci.";
RL J. BACTERIOL. 179:5632-5635(1997).
DR EMBL; U94707; G2149907; -.
KW CELL DIVISION.
SQ SEQUENCE 385 AA; 43741 MW; EF60EE53 CRC32;

Query Match 20.7%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 268 SKTNKD 273
    |||||
QY 5 SKTNKD 10

```

```

RESULT 13
ID P73474 PRELIMINARY; PRT: 498 AA.
AC P73474;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 54.5 KD PROTEIN.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-PCC6803;
RA TABATA S.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-PCC6803;
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUO K., OKUMURA S.,
RA SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. PCC6803. II. Sequence determination of the entire
RT genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
DR EMBL; D90906; D1018247; -.
DR PFAM; PF00916; Sulfate_transp; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 498 AA; 54495 MW; 7FC21892 CRC32;

Query Match 20.7%; Score 6; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 173 FPSPLV 178
    |||||
QY 14 FPSPLV 19

```

```

RESULT 14
ID Q66721 PRELIMINARY; PRT: 512 AA.
AC Q66721;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE (GAG) GENE (POL) (FRAGMENT).
GN GAG.
OS EQUINE INFECTIOUS ANEMIA VIRUS (EIAV).
OC VIRUSES; RETROID VIRUSES; RETROVIRIDAE; LENTIVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-WSU5;
RX MEDLINE; 94201764.
RA MCGUIRE T.C., O'ROURKE K.I., BASZLER T.O., LEIB S.R., BRASSFIELD A.L.,
RA DAVIS W.C.;
RT "Expression of functional protease and subviral particles by vaccinia
RT virus containing equine infectious anaemia virus gag and 5' pol
RT genes.";
RL J. GEN. VIROL. 75:895-900(1994).
DR EMBL; L06609; G323775; -.
DR PFAM; PF00098; zf-CCHC; 2.
DR PFAM; PF00507; gag_p24; 1.
FT NON_TER 1
SQ SEQUENCE 512 AA; 57992 MW; B3169AD2 CRC32;

Query Match 20.7%; Score 6; DB 14; Length 512;
Best Local Similarity 100.0%; Pred. No. 5.03e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 302 NIRQGA 307

```

Qy 23 NIRQGA 28  
|||||  
RESULT 15  
ID Q25397 PRELIMINARY; PRT; 520 AA.  
AC Q25397;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE CONOPRESSIN RECEPTOR 2.  
GN LSCPR2.  
OS LYNNAEA STAGNALIS (GREAT POND SNAIL).  
OC EUKARYOTA; METAZOA; MOLLUSCA; GASTROPODA; PULMONATA; BASOMMATOPHORA;  
OC LYNNAEIDAE; LYNNAEA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=VAS DEFERENS;  
RX MEDLINE: 96216462.  
RA VAN KSTEREN R.E., TENSEN C.P., SMIT A.B., VAN MINNEN J.,  
RA KOLAKOWSKI L.F., MEYERHOF W., RICHTER D., HEERIKHUIZEN H.,  
RA VREUGDENHIL E., GERAERTS W.P.M.;  
RT "Co-evolution of ligand-receptor pairs in the vasopressin/oxytocin  
superfamily of bioactive peptides,"  
RL J. BIOL. CHEM. 271:3619-3626(1996).  
DR EMBL: U40491; G1145726; -.  
DR PFAM: PF00001; 7um\_1; 1.  
SQ SEQUENCE 520 AA; 59435 MW; A01248D8 CRC32;  
  
Query Match 20.7%; Score 6; DB 5; Length 520;  
Best Local Similarity 100.0%; Pred. No. 5.03e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 469 TVTSKT 474  
Qy 2 TVTSKT 7  
|||||

Search completed: Sat Aug 28 14:57:48 1999  
Job time : 50 secs.



\*\*\*\*\*

WQREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:21:34 1999; MasPar time 15.07 Seconds  
Tabular output not generated. 22.584 Million cell updates/sec

Title: >US-09-049-696-49  
Description: (1-16) from US09049696.pep  
Perfect Score: 118  
Sequence: 1 EINKDDVQHKQVCFSR 16

Scoring table: PAM 150  
Gap 15

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 18.400; Variance 54.185; scale 0.340

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	118	100.0	228	20	Human colon specific	1.09e-06
2	118	100.0	228	29	Protein sequence enco	1.09e-06
3	61	51.7	640	22	H. pylori cell envelo	9.94e+00
4	58	49.2	1009	35	Murine protein tyrosi	2.15e+01
5	58	49.2	1009	17	Protein tyrosine kinas	2.15e+01
6	58	49.2	1009	32	Human related adhesio	2.15e+01
7	58	49.2	1009	32	Human PK-2 protein.	2.15e+01
8	58	49.2	1009	35	Mouse related adhesio	2.15e+01
9	58	49.2	1009	35	Mouse protein tyrosin	2.15e+01
10	56	47.5	473	33	Escherichia coli treh	3.56e+01
11	56	47.5	476	20	E. coli trehalose pho	3.56e+01
12	56	47.5	476	19	E. coli trehalose pho	3.56e+01
13	56	47.5	476	13	E. coli trehalose pho	3.56e+01
14	56	47.5	476	28	E. coli trehalose pho	3.56e+01
15	55	46.6	672	39	MEKK1 protein.	4.58e+01
16	55	46.6	672	12	Mammalian MEK kinase	4.58e+01

17	55	46.6	672	15	R77544	MEKK1 protein.	4.58e+01
18	55	46.6	672	30	W56157	A murine mitogen-acti	4.58e+01
19	54	45.8	2842	12	R63508	Adenomatous polyposis	5.87e+01
20	54	45.8	2843	26	W35392	Human adenomatous pol	5.87e+01
21	54	45.8	2843	12	R58634	Adenomatous polyposis	5.87e+01
22	54	45.8	2843	21	W11922	Adenomatous polyposis	5.87e+01
23	54	45.8	2843	35	W76140	Human APC protein #1.	5.87e+01
24	54	45.8	2843	5	R26052	APC gene product in f	5.87e+01
25	54	45.8	2843	35	W76144	Human APC protein #2	5.87e+01
26	54	45.8	2843	27	W38370	Human adenomatous pol	5.87e+01
27	54	45.8	2860	12	R63507	Adenomatous polyposis	5.87e+01
28	54	45.8	2973	37	W76821	Human APC protein.	5.87e+01
29	52	44.1	332	17	R88353	Mutant A-53-Q fragmen	9.62e+01
30	52	44.1	365	31	W51252	G-protein coupled rec	9.62e+01
31	52	44.1	375	31	W51253	G-protein coupled rec	9.62e+01
32	52	44.1	531	31	W51251	G-protein coupled rec	9.62e+01
33	51	43.2	142	8	R47223	Charcot-Leyden Crysta	1.33e+02
34	50	42.4	156	38	W70585	Rat interferon-gamma.	1.57e+02
35	50	42.4	164	13	R66945	Deduced sequence of h	1.57e+02
36	50	42.4	164	23	W24470	Human methenyltetrahy	1.57e+02
37	50	42.4	290	35	W79398	Staphylococcus aureus	1.57e+02
38	50	42.4	344	5	R03989	Lambda-SM70 encoding	1.57e+02
39	49	41.5	186	36	W80402	A secreted protein en	1.99e+02
40	49	41.5	278	3	P60007	Sequence of oncogene	1.99e+02
41	49	41.5	499	27	W37508	Human GSPT1-TK.	1.99e+02
42	48	40.7	205	26	W29224	Nitrile hydratase alp	2.53e+02
43	48	40.7	1040	39	W73584	RAG-1 protein.	2.53e+02
44	48	40.7	1470	28	W23411	Porcine transmissible	2.53e+02
45	47	39.8	585	22	W14915	Modified glutamic aci	3.21e+02

ALIGNMENTS

RESULT 1  
ID W06548 standard; Protein; 228 AA.  
AC W06548;  
DT 13-MAR-1997 (first entry)  
DE Human colon specific gene CSG5 polypeptide fragment.  
KW Colon specific gene; CSG5; colon cancer; metastasis; diagnosis;  
KW therapy; antibody; vaccine.  
OS Homo sapiens.  
PN W09639419-A1.  
PD 12-DEC-1996.  
PF 06-JUN-1995; U07289.  
PR 06-JUN-1995; WO-U07289.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Yu G;  
DR WPI: 97-043054/04.  
DR N-PSDB: T45884.  
PT Human colon specific genes and their expression products - detection  
PT of which, in non-colon tissue samples, can be used as indication of  
PT colon cancer metastasis  
PS Claim 8; Fig 5; 60pp; English.  
CC Novel polypeptides (W06545-53) are encoded by cDNA clones (see also  
CC T45880-92) corresponding to 13 human colon specific genes,  
CC designated CSG1, CSG2, etc., that are primarily expressed in  
CC tissues derived from the colon. Recombinant CSG polypeptides can  
CC be produced in transformed host cells. They are useful diagnostic  
CC markers for colon cancer and for colon cancer metastasis and can  
CC also be used to screen for (ant)agonist cpds. of therapeutic or  
CC diagnostic value. Antibodies raised against the colon-specific  
CC polypeptides may be used to target colon cancer cells or as part  
CC of a colon cancer vaccine.  
SQ Sequence 228 AA;

Query Match 100.0%; Score 118; DB 20; Length 228;

Best Local Similarity 100.0%; Pred.No. 1.09e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 59 einkddvqhkqvcfsr 74

|||||

QY 1 EINKDDVQHKQVCFSR 16

RESULT 2  
ID W46879 standard; Protein: 228 AA.  
AC W46879;  
DT 22-JUN-1998 (first entry)  
DE Protein sequence encoded by a colon-specific gene.  
KW Colon-specific gene; probe; detection; expression; human;  
KW diagnostic assay; colon cancer; antibody; screening.  
OS Homo sapiens.  
PN US5733748-A.  
PD 31-MAR-1998.  
PF 06-JUN-1995; 459667.  
PR 08-JUN-1995; US-469667.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen C, Yu G;  
DR WPI: 98-229823/20.  
DR N-PSDB: V16672.  
DT Colon-specific nucleic acids - useful as probes for detecting colon cancer micrometastases  
PS Claim 1; Fig 5; 51pp; English.  
CC W46876-80 and W4682-85 represent proteins encoded by colon-specific genes. The polynucleotides encoding these proteins can be used as probes to detect expression of the corresponding human genes, e.g. in diagnostic assays for detecting micrometastases of colon cancer. Recombinant cells containing the polynucleotides can be used to produce the proteins, in order that antibodies can be raised and used in further screening or diagnostics.  
CC used in further screening or diagnostics.  
SQ Sequence 228 AA;  
Query Match 100.08; Score 118; DB 29; Length 228;  
Best Local Similarity 100.08; Pred. No. 1.09e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 59 einkddvghkvqcf 74  
QY 1 EINKDDVGHKVQCF 16  
|||||||  
RESULT 3  
ID W20955 standard; Protein: 640 AA.  
AC W20955;  
DT 21-JUL-1997 (first entry)  
DE H. pylori cell envelope inner membrane protein hplp1392orf22.  
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
OS Helicobacter pylori.  
PN W09640893-A1.  
PD 19-DEC-1996.  
PF 06-JUN-1996; U09122.  
PR 07-JUN-1995; US-487032.  
PR 01-APR-1996; US-630405.  
PA (ASTR.) ASTRA AB.  
PI Berglindh OT, Smith D, Mellgaerd BL;  
DR WPI: 97-052306/05.  
DR N-PSDB: T68208.  
DT Helicobacter pylori nucleic acid sequences and related polypeptide(s) - useful for vaccines to treat or prevent H. pylori infection, and to detect Helicobacter  
PT infection, and to detect Helicobacter  
PS Claim 56; Page 1346-1348; 1481pp; English.  
CC The present sequence represents a Helicobacter pylori cell envelope inner membrane protein.  
CC The protein may be used in a vaccine to prevent or treat H. pylori infection or to identify H. pylori polypeptide binding compounds, useful as potential H. pylori life cycle activators or inhibitors.  
CC The genomic sequence of H. pylori (AtCC 55679) was determined from overlapping contigs generated by mechanically shearing the bacterial DNA. The sequences were analysed for ORF of at least 180 nucleotides, and the predicted coding regions defined by computer evaluation. To identify likely H. pylori antigens for vaccine development, the amino acid sequences predicted from various ORF were analysed for significant homology to other known or exported membrane proteins. Having identified and determined the sequences of interest, particular regions can be

CC isolated from H. pylori by PCR amplification for recombinant polypeptide production, e.g. in E. coli hosts.  
SQ Sequence 640 AA;

Query Match 51.7%; Score 61; DB 22; Length 640;  
Best Local Similarity 56.7%; Pred. No. 9.94e+00;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 163 dinhkevfc 171  
QY 6 DVQHKQVCF 14  
|: |||||

RESULT 4  
ID W64568 standard; Protein: 1009 AA.  
AC W64568;  
DT 24-NOV-1998 (first entry)  
DE Murine protein tyrosine kinase 2 (PYK2).  
KW Protein tyrosine kinase 2; PYK2 gene; mouse; podosome;  
KW cell adhesion kinase beta; related adhesion focal tyrosine kinase;  
KW focal adhesion kinase; ligand; monocyte; osteoporosis;  
KW inflammation; therapy.  
OS Mus sp.  
FH Key Location/Qualifiers  
FT Domain 419..679 /note= "kinase domain"  
FT Domain 701..767  
FT /note= "proline-rich domain"  
PN W09835056-A1.  
PD 13-AUG-1998.  
PF 09-FEB-1998; U02797.  
PR 11-FEB-1997; US-037560.  
PA (MERI) MERCK & CO INC.  
PI Duong Le T, Rodan GA;  
DR WPI: 98-447250/38.  
DR N-PSDB: V49608.  
PT Identifying agents that bind and modulate protein tyrosine kinase 2 - useful for inhibiting migration, adhesion or activity of monocytic cells, particularly for treatment and prevention of osteoporosis and inflammation  
PT Inflammation  
PS Example 3; Fig 8; 56pp; English.  
CC This is the amino acid sequence of murine protein tyrosine kinase 2 (PYK2), deduced from a cDNA sequence (see V49608) derived from 2 overlapping clones of a mouse spleen library. PYK2, also named cell adhesion kinase beta and related adhesion focal tyrosine kinase, is a recently described member of the focal adhesion kinase family. A claimed method for identifying compounds (I) that bind to, and/or modulate, PYK2 activity comprises treating PYK2 with a test compound, then determining binding and optionally any change in activity. The method is used to identify (I) that prevent adhesion of monocytes to substrate, inhibit osteoclast mobility or prevent monocytic cells from degrading extracellular matrix. These are potentially useful for treatment and prevention of osteoporosis and inflammation, or generally any condition associated with the migration and invasion of monocytes. The method is based on the observation that PYK2 is primarily responsible for podosome formation in monocytes and macrophages.  
CC formation in monocytes and macrophages.  
SQ Sequence 1009 AA;

Query Match 49.2%; Score 58; DB 35; Length 1009;  
Best Local Similarity 42.9%; Pred. No. 2.15e-01;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 dvekedvrlkvfc 45  
QY 1 EINKDDVGHKVQCF 14  
|: |||||

RESULT 5  
ID R98351 standard; Protein: 1009 AA.  
AC R98351;  
DT 29-AUG-1996 (first entry)  
DE Probin tyrosine kinase PYK2.

KW Probin tyrosine kinase: PYK2; signal transduction: ion channel;  
KW probe: diagnosis; gene therapy; epilepsy; schizophrenia;  
KW hyperactivity; pain; stroke; Alzheimer disease; migraine;  
KW Parkinson disease.  
OS Homo sapiens.

FT Key Location/Qualifiers  
FT modified\_site 402  
FT /note= "predicted autophosphorylation site that  
FT may be involved in src SH2 domain  
FT binding"  
FT 779..882  
FT binding\_site  
FT /note= "predicted binding site for GRB2 SH2 domain"  
FT binding\_site 904..907  
FT /note= "predicted binding site for SHP2"  
FT WO9618738-A2.  
PN 20-JUN-1996.  
PD 06-DEC-1995; UI5846.  
PF 15-DEC-1994; US-357642.  
PR 02-JUN-1995; US-460626.  
PR (SUGEN) SUGEN INC.  
PA (UYN) UNIV NEW YORK STATE.  
PI Lev S, Schlessinger J;  
PI WPI: 96-300655/30.  
DR N-PSDB: T30373.  
PT Probin tyrosine kinase 2 and nucleic acid encoding PYK2 - used to  
PT detect and treat diseases characterised by a signal transduction  
PT pathway abnormality, e.g. epilepsy or schizophrenia  
PS Example 1: Page 123-125; 140pp; English.  
CC A novel non-receptor tyrosine kinase (R98351), designated probin  
CC tyrosine kinase 2 (PYK2), was identified as the product of a  
CC cDNA clone (T30373) derived from human brain. The 111 kDa PYK2  
CC is activated by binding of ligand to G-coupled protein receptors  
CC and is believed to regulate the activity of potassium channels in  
CC response to neurotransmitter signalling. Recombinant PYK2 can  
CC be produced in host cells. It can be used to detect and treat  
CC conditions characterised by an abnormality in signal transduction,  
CC e.g. epilepsy, schizophrenia, hyperactivity, pain, stroke, migraine  
CC Alzheimer disease and Parkinson disease. It may also be expressed  
CC in vivo for gene therapy.  
SQ Sequence 1009 AA;

Query Match 49.2%; Score 58; DB 17; Length 1009;  
Best Local Similarity 42.9%; Pred. No. 2.15e+01;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 dvekdvrlkvkvcf 45

QY 1 EINKDDVQHKKVCF 14

RESULT 6  
ID W48373 standard; Protein; 1009 AA.

AC W48373;  
DE 17-AUG-1998 (first entry)  
DE Human related adhesion focal tyrosine kinase (RAFTK).  
KW RAFTK; related adhesion focal tyrosine kinase; human;  
KW protein tyrosine kinase; PTK; cell growth; cell differentiation;  
KW cell adhesion; cell migration; regulation; signal transduction;  
KW metastasis; cancer; leukaemia; thrombocytopaenia; therapy.  
OS Homo sapiens.

FT Key Location/Qualifiers  
FT Domain 1..39  
FT /note= "N-terminal domain"  
FT Domain 427..679  
FT /label= kinase  
FT /note= "domain"  
FT Domain 690..767  
FT /note= "C-terminal domain"  
FT Binding\_site 432..437  
FT /note= "ATP binding site"  
FT Binding\_site 402  
FT /note= "residue predicted to interact with the  
FT gamma phosphate group of bound ATP"

FT Binding\_site 529  
FT /note= "residue predicted to interact with the  
FT gamma phosphate group of bound ATP"  
FT Binding\_site 655  
FT /note= "residue predicted to interact with the  
FT gamma phosphate group of bound ATP"  
FT Misc\_difference 549..554  
FT /note= "PTK conserved region"  
FT Misc\_difference 588..592  
FT /note= "PTK conserved region"  
FT Region 880..887  
FT /note= "region important in mediating association  
FT of RAFTK with src and/or fyn, particularly  
FT when Tyr-881 is phosphorylated"  
FT Modified\_site 881  
FT /note= "phosphorylated"  
FT WO9807870-A1.  
PN 26-FEB-1998.  
PD 12-AUG-1997; UI4093.  
PF 13-MAR-1997; US-816462.  
PR 23-AUG-1996; US-703623.  
PR (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
PI Avraham H, Avraham S, Groopman JE;  
PI WPI: 98-169174/15.  
DR N-PSDB: V17800.

PT New nucleic acid encoding vertebrate related adhesion focal tyrosine  
PT kinase - useful for, e.g. regulating growth, differentiation,  
PT adhesion and migration of cells, used for treating metastases,  
PS Claim 5: Page 136-139; 167pp; English.  
CC This protein comprises novel human related adhesion focal tyrosine  
CC kinase (RAFTK), an intracytoplasmic protein tyrosine kinase (PTK)  
CC involved in the modulation of cell growth and differentiation. Its  
CC amino acid sequence was deduced from cDNA clones (see V17800)  
CC obtained from a hippocampal library. RAFTK mRNA is expressed most  
CC highly in the brain, especially the amygdala and hippocampus, and  
CC also in megakaryocytes. RAFTK may be expressed in host cells or  
CC in transgenic animals and used to raise antibodies and to screen  
CC for compounds capable of modulating RAFTK interaction with cellular  
CC proteins. Agents that modulate RAFTK are used to control growth,  
CC differentiation, haematopoiesis or survival of cells (claimed)  
CC (especially mast cells, melanocytes and megakaryocytes), or  
CC adhesion (especially focal adhesion formation), migration,  
CC phagocytosis or motility of cells, particularly for treating  
CC metastases. Typical applications are in treatment of immune  
CC mediated diseases, e.g. cancer or leukaemia, where apoptosis is  
CC induced, inhibition of apoptosis where this is induced by  
CC radiation, chemotherapy or neurodegenerative disease, and also for  
CC expanding cells for transplantation, treating conditions  
CC associated with megakaryocyte abnormality, e.g. thrombocytopaenia,  
CC and control of platelet aggregation. RAFTK proteins and modulators  
CC may be expressed from nucleic acids introduced in gene therapy  
CC vectors.  
SQ Sequence 1009 AA;

Query Match 49.2%; Score 58; DB 31; Length 1009;

Best Local Similarity 42.9%; Pred. No. 2.15e+01;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 dvekdvrlkvkvcf 45

QY 1 EINKDDVQHKKVCF 14

RESULT 7

ID W57891 standard; Protein; 1009 AA.  
AC W57891;  
DE 25-SEP-1998 (first entry)  
DE Human PYK-2 protein.  
KW Human; PYK-2; non-receptor tyrosine kinase; indolignone; epilepsy;  
KW schizophrenia; hyperactivity; Alzheimer's disease; Parkinson's disease;  
KW myasthenia gravis; neuroblastoma; snake venom poisoning; therapy;  
KW spinal cord injury; Tourette's syndrome.  
OS Homo sapiens.

```

Query Match          49.2%; Score 58; DB 35; Length 1009;
Best Local Similarity 42.9%; Pred. No. 2.15e+01;
Matches             6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db      32 dvekdvrlkvcf 45
      ::::|::|::|
QY      1 EINKDDVQHKQVCF 14

RESULT 10
ID      W49030 standard; protein: 473 AA.
AC      W49030;
DT      07-OCT-1998 (first entry)

```

DE Escherichia coli trehalose synthase.  
 KW Trehalose synthase; stabilising agent; prokaryotic cell preservation;  
 KW prophylaxis; live bacterial vaccine; vaginal infection;  
 KW urinary tract infection; live bacterial neutraceutical.  
 OS Escherichia coli.  
 PN WO9824882-A1.  
 PS 11-JUN-1998.  
 PD 05-DEC-1997; G03375.  
 PR 05-DEC-1996; US-032423.  
 PA (QUAD-) QUADRANT HOLDINGS CAMBRIDGE LTD.  
 PI Colaco C, Dhaliwal KS, Roser BV, Tunnaciffe AG,  
 PI Welsh DT;  
 DR WPI: 98-333305/29.  
 PT Preservation of prokaryotic cells - by increasing intracellular  
 PT trehalose concentration, mixing with a stabilising agent, and drying  
 PT to produce a glass form of the stabilising agent  
 PS Disclosure: Pages 47-48; 64pp; English.  
 CC The present sequence represents the Escherichia coli trehalose  
 CC synthase. The invention provides methods of preserving prokaryotic  
 CC cells which involves inducing the intracellular trehalose production  
 CC of the prokaryotic cell to an amount which effectively allows increased  
 CC storage stability and drying of the cells in the presence of a  
 CC stabilising agent. Therefore, prokaryotic cells known to synthesise  
 CC trehalose, such as E. coli, can be used to illustrate the methods  
 CC of the invention. Bacteria stabilised by these methods showed less than  
 CC 10% loss of viability on storage even after being stored at temperatures  
 CC up to at least 37 degrees centigrade for as long as 6 weeks. The  
 CC invention claims the methods to be useful for producing dried, stable  
 CC prokaryotic cells such as bacteria useful for pharmacological treatment,  
 CC prophylaxis, agricultural and industrial applications. The invention  
 CC also claims the methods to be useful for producing live bacterial  
 CC vaccines in a dry stable form, live bacterial neutraceuticals in a dry  
 CC stable form and other live bacterial pharmaceuticals in a dry stable  
 CC form, e.g. for treatment of vaginal or urinary tract infections.  
 SQ Sequence 473 AA;

Query Match 47.5%; Score 56; DB 33; Length 473;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 434 ivkndinhwqecf 446  
 QY 2 INKDDVQHKQVCF 14  
 | | | | | | | |

RESULT 11  
 ID W11493 standard; Protein; 476 AA.  
 AC W11493;  
 DT 03-APR-1997 (first entry)  
 DE E. coli trehalose phosphate synthase enzyme.  
 KW Trehalose phosphate synthase; TPS; D-glucosyl D-glucoside; potato;  
 KW disaccharide; recombinant; transgenic plant; food crop; carbohydrate;  
 KW metabolism; animal feed; otsa.  
 OS Escherichia coli clone 7F11.  
 PN WO9501446-A1.  
 PD 12-JAN-1995.  
 PF 30-JUN-1994; E02167.  
 PR 30-JUN-1993; EP-201904.  
 PR 24-AUG-1993; WO-E02290.  
 PA (MOGE-) MOGEN INT NV.  
 PI Does MP, Hoekema A, Pen J, Van Den Elzen PJM;  
 DR WPI: 95-061010/08.  
 DR N-PSDB: T51412.  
 PT Plant expressible gene, encoding trehalose - used to increase  
 PT trehalose production capacity of plants  
 PS Claim 39: Page 33-35; 62pp; English.  
 CC W11493 represents an E. coli-derived trehalose phosphate synthase (TPS)  
 CC enzyme. Trehalose is a general name given to D-glucosyl D-glucoside  
 CC disaccharides based upon alpha-alpha, alpha-beta or beta-beta linked  
 CC glucose moieties. Trehalose is a widespread naturally occurring  
 CC disaccharide. The otsa gene fragment (encoding TPS, see T51412) may be  
 CC used to transform plants to express increased amounts of TPS, thus  
 CC increasing the amount of trehalose produced by the plant. However the

CC substrates (UDPG and F6P) used for trehalose production are also used  
 CC in the production of sucrose and ultimately starch. To reduce  
 CC competition for these substrates antisense molecules can be  
 CC administered to the plant which inhibit the expression of other enzymes,  
 CC i.e. sucrose phosphate synthase (see T51421) and AGPase (ADP-Glucose  
 CC phosphorylase), which are involved in the production of sucrose and  
 CC starch, hence increasing the amount of substrate available for TPS and  
 CC increasing trehalose production. Food crops and animal feed plants are  
 CC especially suitable for transformation with a TPS encoding sequence  
 CC to produce plants having a high "simple-carbohydrate" level, i.e.  
 CC high energy foods. Potatoes, turnips, tomatoes, cabbages, yam, wheat,  
 CC corn, soybean and many other plants are listed as suitable in the  
 CC specification.  
 SQ Sequence 476 AA;

Query Match 47.5%; Score 56; DB 20; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 437 ivkndinhwqecf 449  
 QY 2 INKDDVQHKQVCF 14  
 | | | | | | | |

RESULT 12  
 ID R98601 standard; Protein; 476 AA.  
 AC R98601;  
 DT 06-DEC-1996 (first entry)  
 DE E. coli trehalose phosphate synthase.  
 KW Trehalose phosphate synthase; trehalase inhibitor;  
 KW transgenic plant; antisense; potato; Solanum tuberosum; tobacco;  
 KW Nicotiana tabacum.  
 OS Escherichia coli clone 7F11.  
 PN WO9621030-A1.  
 PD 11-JUL-1996.  
 PF 03-JAN-1996; E00080.  
 PR 04-JAN-1995; EP-200008.  
 PR 07-SEP-1995; EP-202415.  
 PA (MOGE-) MOGEN INT NV.  
 PI Goddijn OJM, Krutwagen RWHH, Verwoerd TC, Voogd E;  
 DR WPI: 96-334005/33.  
 DR N-PSDB: T33143.  
 PT Increasing trehalose prodn. in plants by growing in presence of  
 PT trehalase inhibitor - and chimeric gene expressible in plants  
 PT encoding trehalose phosphate synthase  
 PS Disclosure: Page 29-31; 58pp; English.  
 CC Escherichia coli trehalose phosphate synthase (R98601) can be  
 CC expressed in plants, or specific areas of plants such as potato  
 CC tubers, by placing the otsa gene (T33143) under control of e.g.  
 CC the patatin or CAMV 35S RNA promoter, incorporating the construct  
 CC into a vector, and transforming plant cells utilising Agrobacterium  
 CC tumefaciens. Accumulation of trehalose in the transgenic plants is  
 CC further enhanced by the presence of a trehalase inhibitor, such as  
 CC the cockroach (Periplaneta americana) 86 kDa trehalase inhibitor,  
 CC validamycin C, or an antisense trehalase gene (see also T33142).  
 CC Trehalose protects the plant against drought, and may be extracted  
 CC for industrial use.  
 SQ Sequence 476 AA;

Query Match 47.5%; Score 56; DB 19; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 437 ivkndinhwqecf 449  
 QY 2 INKDDVQHKQVCF 14  
 | | | | | | | |

RESULT 13  
 ID R71412 standard; Protein; 476 AA.  
 AC R71412;  
 DT 03-OCT-1995 (first entry)  
 DE E. coli trehalose phosphate synthase.

KW Plant; trehalose; carbohydrate.  
 OS Escherichia coli.  
 PN WO9506126-A.  
 PD 02-MAR-1995.  
 PF 24-AUG-1993; E02290.  
 PR 24-AUG-1993; WO-E02290.  
 PA (MOGE-) MOGEN INT NV.  
 PI Does MP, Hoekema A, Pen J, Van DEN ELZEN PJM;  
 DR WPI: 95-106855/14.  
 N-PSDB: Q84949.  
 PT Plant expressible gene increasing the trehalose content of  
 PT transformed plants - inhibits endogenous genes to enhance  
 PT substrate availability for trehalose phosphate synthase  
 PS Disclosure: Page 39; 75pp; English.  
 CC The sequence is that of E. coli trehalose phosphate from the gene,  
 CC ostaA. The gene is used to transform plant cells to produce plants  
 CC and plant cell cultures capable of producing trehalose. Such  
 CC transformed plants may be used to extract specific carbohydrates, or  
 CC they may be processed as food having improved properties due to the  
 CC presence of the carbohydrate cpds.  
 SQ Sequence 476 AA;  
 Query Match 47.5%; Score 56; DB 13; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Db 437 ivkndinhwqecf 449  
 | | | | | | | |  
 QY 2 INKDDVQHKKVCF 14  
 RESULT 14  
 ID W41817 standard; Protein; 476 AA.  
 AC W41817;  
 DT 14-MAY-1998 (first entry)  
 DE E. coli trehalose phosphate synthase partial protein.  
 KW Trehalose-6-phosphate; trehalase; trehalose phosphate synthase;  
 KW trehalose phosphate phosphatase; trehalase; treatment; tumour;  
 KW diabetes; obesity; transgenic animal.  
 OS Escherichia coli.  
 PN WO974326-A2.  
 PD 13-NOV-1997.  
 PF 02-MAY-1997; E02497.  
 PR 29-AUG-1996; EP-202395.  
 PR 03-MAY-1996; EP-201225.  
 PR 26-JUL-1996; EP-202128.  
 PA (MOGE-) MOGEN INT NV.  
 PI Goddijn OJM, Pen J, Smeekens JCM, Smits MT;  
 DR WPI: 97-558986/51.  
 N-PSDB: VI3269.  
 PT Modification of trehalose-6-phosphate levels - for development and,  
 PT or composition of cells, tissues and organs in plants  
 PS Disclosure: Pages 79-81; 193pp; English.  
 CC The present sequence was used in the development of a novel method  
 CC for the modification of the development and/or composition of  
 CC cells, tissues or organs in vivo by inducing a change in the  
 CC metabolic availability of trehalose-6-phosphate (T6P) or trehalase.  
 CC Increasing the availability of T6P inhibits carbon flow in the  
 CC glycolytic direction, stimulates photosynthesis, stimulates sink  
 CC related activity and inhibits growth of a cell or tissue.  
 CC specifically in plants. Decreasing availability of T6P stimulates  
 CC carbon flow in the glycolytic direction, inhibits photosynthesis,  
 CC stimulates cell or tissue growth and increases metabolism of cells.  
 CC Trehalose phosphate synthase (TPS) is capable of forming T6P and  
 CC trehalase phosphate phosphatase (TPP) degrades T6P into trehalose.  
 CC Therefore TPS and TPP can be expressed in cells to change the  
 CC availability of T6P to induce the above changes. TPS and TPP can  
 CC also be used to treat, e.g. tumours, diabetes or obesity, or in  
 CC transgenic animals to alter protein/carbohydrate balance.  
 SQ Sequence 476 AA;  
 Query Match 47.5%; Score 56; DB 28; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Db 437 ivkndinhwqecf 449  
 | | | | | | | |  
 QY 2 INKDDVQHKKVCF 14  
 RESULT 14  
 ID W41817 standard; Protein; 476 AA.  
 AC W41817;  
 DT 14-MAY-1998 (first entry)  
 DE E. coli trehalose phosphate synthase partial protein.  
 KW Trehalose-6-phosphate; trehalase; trehalose phosphate synthase;  
 KW trehalose phosphate phosphatase; trehalase; treatment; tumour;  
 KW diabetes; obesity; transgenic animal.  
 OS Escherichia coli.  
 PN WO974326-A2.  
 PD 13-NOV-1997.  
 PF 02-MAY-1997; E02497.  
 PR 29-AUG-1996; EP-202395.  
 PR 03-MAY-1996; EP-201225.  
 PR 26-JUL-1996; EP-202128.  
 PA (MOGE-) MOGEN INT NV.  
 PI Goddijn OJM, Pen J, Smeekens JCM, Smits MT;  
 DR WPI: 97-558986/51.  
 N-PSDB: VI3269.  
 PT Modification of trehalose-6-phosphate levels - for development and,  
 PT or composition of cells, tissues and organs in plants  
 PS Disclosure: Pages 79-81; 193pp; English.  
 CC The present sequence was used in the development of a novel method  
 CC for the modification of the development and/or composition of  
 CC cells, tissues or organs in vivo by inducing a change in the  
 CC metabolic availability of trehalose-6-phosphate (T6P) or trehalase.  
 CC Increasing the availability of T6P inhibits carbon flow in the  
 CC glycolytic direction, stimulates photosynthesis, stimulates sink  
 CC related activity and inhibits growth of a cell or tissue.  
 CC specifically in plants. Decreasing availability of T6P stimulates  
 CC carbon flow in the glycolytic direction, inhibits photosynthesis,  
 CC stimulates cell or tissue growth and increases metabolism of cells.  
 CC Trehalose phosphate synthase (TPS) is capable of forming T6P and  
 CC trehalase phosphate phosphatase (TPP) degrades T6P into trehalose.  
 CC Therefore TPS and TPP can be expressed in cells to change the  
 CC availability of T6P to induce the above changes. TPS and TPP can  
 CC also be used to treat, e.g. tumours, diabetes or obesity, or in  
 CC transgenic animals to alter protein/carbohydrate balance.  
 SQ Sequence 476 AA;  
 Query Match 47.5%; Score 56; DB 28; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Db 437 ivkndinhwqecf 449  
 | | | | | | | |  
 QY 2 INKDDVQHKKVCF 14  
 RESULT 14  
 ID W41817 standard; Protein; 476 AA.  
 AC W41817;  
 DT 14-MAY-1998 (first entry)  
 DE E. coli trehalose phosphate synthase partial protein.  
 KW Trehalose-6-phosphate; trehalase; trehalose phosphate synthase;  
 KW trehalose phosphate phosphatase; trehalase; treatment; tumour;  
 KW diabetes; obesity; transgenic animal.  
 OS Escherichia coli.  
 PN WO974326-A2.  
 PD 13-NOV-1997.  
 PF 02-MAY-1997; E02497.  
 PR 29-AUG-1996; EP-202395.  
 PR 03-MAY-1996; EP-201225.  
 PR 26-JUL-1996; EP-202128.  
 PA (MOGE-) MOGEN INT NV.  
 PI Goddijn OJM, Pen J, Smeekens JCM, Smits MT;  
 DR WPI: 97-558986/51.  
 N-PSDB: VI3269.  
 PT Modification of trehalose-6-phosphate levels - for development and,  
 PT or composition of cells, tissues and organs in plants  
 PS Disclosure: Pages 79-81; 193pp; English.  
 CC The present sequence was used in the development of a novel method  
 CC for the modification of the development and/or composition of  
 CC cells, tissues or organs in vivo by inducing a change in the  
 CC metabolic availability of trehalose-6-phosphate (T6P) or trehalase.  
 CC Increasing the availability of T6P inhibits carbon flow in the  
 CC glycolytic direction, stimulates photosynthesis, stimulates sink  
 CC related activity and inhibits growth of a cell or tissue.  
 CC specifically in plants. Decreasing availability of T6P stimulates  
 CC carbon flow in the glycolytic direction, inhibits photosynthesis,  
 CC stimulates cell or tissue growth and increases metabolism of cells.  
 CC Trehalose phosphate synthase (TPS) is capable of forming T6P and  
 CC trehalase phosphate phosphatase (TPP) degrades T6P into trehalose.  
 CC Therefore TPS and TPP can be expressed in cells to change the  
 CC availability of T6P to induce the above changes. TPS and TPP can  
 CC also be used to treat, e.g. tumours, diabetes or obesity, or in  
 CC transgenic animals to alter protein/carbohydrate balance.  
 SQ Sequence 476 AA;  
 Query Match 47.5%; Score 56; DB 28; Length 476;  
 Best Local Similarity 53.8%; Pred. No. 3.56e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Db 437 ivkndinhwqecf 449  
 | | | | | | | |  
 QY 2 INKDDVQHKKVCF 14  
 RESULT 15  
 ID W73536 standard; Protein; 672 AA.  
 AC W73536;  
 DT 04-MAR-1999 (first entry)  
 DE MEKK1 protein.  
 KW Mitogen activated protein kinase kinase kinase; MEKK; MAPK; cancer;  
 KW apoptosis regulator; autoimmune disease; inflammation; allergy; therapy;  
 KW neuronal disorder.  
 OS Mus sp.  
 PN US5854043-A.  
 PD 29-DEC-1998.  
 PF 14-OCT-1994; 323460.  
 PR 14-OCT-1994; US-323460.  
 PR 15-APR-1993; US-049254.  
 PR 15-APR-1994; WO-U04178.  
 PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
 PI JOHNSON GL;  
 DR WPI: 99-094912/08.  
 N-PSDB: V45619.  
 PT Mitogen activated protein kinase kinase kinases and their fragments -  
 PT used for regulating signalling from growth factor receptors, e.g. to  
 PT modulate apoptosis for treatment of cancer, autoimmune disease and  
 PT inflammation  
 PS Example 1; Column 11-16; 96pp; English.  
 CC This sequence is the MEKK2 protein of the invention. MEKK proteins  
 CC are mitogen-activated protein kinase (MAPK) kinase (MEK) kinase proteins.  
 CC MEKKs phosphorylate and activate MEK proteins and other signal  
 CC transduction molecules, so can regulate signalling initiated from a  
 CC growth factor receptor in a way different from that involving Raf  
 CC protein. Particularly MEKKs, or their fragments, are involved in  
 CC regulation of apoptosis so they, or agents that increase their activity,  
 CC are used to treat cancers, autoimmune diseases, inflammation, allergies,  
 CC neuronal disorders (e.g. Alzheimer's or Parkinson's diseases) and in  
 CC wound healing. MEKKs are also useful for identifying agents that  
 CC regulate signal transduction from cell surface receptors (e.g. from their  
 CC effect on ability of MEKK to phosphorylate a substrate such as MEK or Jun  
 CC extracellular signal-regulated kinase).  
 SQ Sequence 672 AA;  
 Query Match 46.6%; Score 55; DB 39; Length 672;  
 Best Local Similarity 53.8%; Pred. No. 4.58e+01;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Db 325 dtykddvnhqkc 337  
 | | | | | | | |  
 QY 1 EINKDDVQHKKVCF 13  
 Search completed: Mon Aug 23 13:21:57 1999  
 Job time : 23 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:20:59 1999; MasPar time 8.30 Seconds

Tabular output not generated. 77.274 Million cell updates/sec

Title: >US-09-049-696-49  
Description: (1-16) from US09049696.ppep  
Perfect Score: 118  
Sequence: 1 EINKDDVOHKQVCFR 16

Scoring table: PAM 150  
Gap 15

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 25.497; Variance 34.347; scale 0.742

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	61	51.7	806	2	A71979	hypothetical protein
2	61	51.7	831	2	B64528	conserved hypothetical protein
3	60	50.8	251	2	A69897	hypothetical protein
4	58	49.2	150	2	B57434	protein-tyrosine kinase
5	58	49.2	1009	2	A57434	protein-tyrosine kinase
6	58	49.2	1009	2	G02330	cell adhesion kinase
7	58	49.2	1009	2	S60238	tyrosine-specific protein
8	57	48.3	913	3	JG0168	gob-5 protein - mouse
9	56	47.5	474	2	I83402	alpha,alpha-trehalose
10	56	47.5	692	1	S46953	phosphotransferase
11	56	47.5	1203	2	A33165	finger protein sdc-1
12	55	46.6	328	2	D41433	probable ABC-type transporter
13	55	46.6	333	2	S30915	proteolipophilin
14	55	46.6	687	2	A46212	MEK kinase - mouse
15	55	46.6	699	2	A57467	transcription factor
16	54	45.8	180	2	S57944	ADP-ribosylation factor
17	54	45.8	180	2	I55371	ADP-ribosylation factor
18	54	45.8	180	2	JC4948	ADP-ribosylation factor
19	54	45.8	180	2	B38622	ADP-ribosylation factor
20	54	45.8	180	2	JC4949	ADP-ribosylation factor
21	54	45.8	180	2	A53859	ADP-ribosylation factor
22	54	45.8	180	2	A23741	ADP-ribosylation factor
23	54	45.8	201	2	A53688	5-formyltetrahydrofol

24	54	45.8	278	2	S60619	homeotic protein cnox
25	54	45.8	328	2	B49543	maltoaccharide-utili
26	54	45.8	1310	2	S30328	multidrug resistance
27	54	45.8	2843	1	RBHUAP	adenomatous polyposis
28	54	45.8	2845	2	I49505	adenomatous polyposis
29	53	44.9	273	2	C71615	protein of the YMR7 f
30	53	44.9	280	2	B71502	probable acyl-carrier
31	53	44.9	354	2	G71465	hypothetical protein
32	53	44.9	395	2	A37350	cyclin B - starfish
33	53	44.9	1078	3	T00494	DNA-dependent RNA pol
34	53	44.9	1407	1	RNECC	DNA-directed RNA pol
35	53	44.9	3660	1	S02041	dystrophin, muscle
36	52	44.1	231	2	S75380	hypothetical protein
37	52	44.1	365	2	S68208	G protein-coupled rec
38	52	44.1	409	2	S00662	cyclin - sea urchin
39	52	44.1	604	2	S15794	lin-9 protein - Caeno
40	52	44.1	892	2	A56213	DNA excision-repair p
41	52	44.1	1358	1	XOCHDH	xanthine dehydrogenas
42	51	43.2	142	2	A46523	Charcot-Leyden crysta
43	51	43.2	273	2	D64645	hypothetical protein
44	51	43.2	480	2	E70042	carboxy-terminal proc
45	51	43.2	1495	2	S22610	DNA (cytosine-5-)met

ALIGNMENTS

RESULT 1  
ENTRY  
TITLE A71979 #type complete  
hypothetical protein jhp0061 - Helicobacter pylori (strain J99)  
ORGANISM #formal\_name Helicobacter pylori  
DATE 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 12-Feb-1999  
ACCESSIONS A71979  
REFERENCE A71800  
#authors Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.  
#journal Nature (1999) 397:176-180  
#title Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.  
#cross-references MUID:99120557  
#accession A71979  
#status preliminary  
#molecule\_type DNA  
#residues 1-806 #label ARN  
#cross-references GB:AE001445; GB:AE001439; MID:g4154559; PID:g4154569  
#experimental\_source strain J99

GENETICS  
#gene jhp0061  
#length 806 #molecular-weight 92770 #checksum 1838

Summary  
Query Match 51.7%; Score 61; DB 2; Length 806;  
Best Local Similarity 66.7%; Pred. No. 4.03e-01;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 329 DINKHEVCF 337  
|: |||||  
QY 6 DVQHKQVCF 14

RESULT 2  
ENTRY  
TITLE B64528 #type complete  
conserved hypothetical ATP-binding protein HP0066 - Helicobacter pylori (strain 26695)  
ORGANISM #formal\_name Helicobacter pylori  
DATE 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 10-Jul-1998  
ACCESSIONS B64528

```

Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Taconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenberg, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenecker, I.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
#cross-references MUID:98044033
#accession A69897
##status preliminary; nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-251 #label KUN
##cross-references GB:299114; GB:AL009126; NID:g2634230; PID:el185341; PID:g2634262
##experimental_source strain 168
GENETICS
#gene yoaP
SUMMARY
#length 251 #molecular-weight 28855 #checksum 8603
Query Match 50.8%; Score 60; DB 2; Length 251;
Best Local Similarity 38.5%; Pred. No. 6.35e-01;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
Db 5 EYKDNIEDNHIC 17
|| |||:: |||
QY 1 EINKDDVQHKQVC 13

RESULT 4
ENTRY B57434 #type fragment
TITLE protein-tyrosine kinase (EC 2.7.1.112) CAK-beta - human (fragment)
ALTERNATE_NAMES cell adhesion kinase-beta
ORGANISM #formal_name Homo sapiens #common_name man
DATE 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 31-Oct-1997
ACCESSIONS B57434
REFERENCE A57434
#authors Sasaki, H.; Nagura, K.; Ishino, M.; Tobioka, H.; Kotani, K.; Sasaki, T.
#journal J. Biol. Chem. (1995) 270:21206-21219
#title Cloning and characterization of cell adhesion kinase beta, a novel protein-tyrosine kinase of the focal adhesion kinase subfamily.
#cross-references MUID:95403356
#accession B57434
##status preliminary
##molecule_type mRNA
##residues 1-150 #label SAS
##cross-references GB:045953; NID:g1008676; PID:g1008884; PID:g1000677
CLASSIFICATION #superfamily unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
KEYWORDS phosphotransferase
SUMMARY
#length 150 #checksum 4845
Query Match 49.2%; Score 58; DB 2; Length 150;
Best Local Similarity 42.9%; Pred. No. 1.56e+00;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Db 32 DVEKEDVRILKVCF 45
:::|::|::|
QY 1 EINKDDVQHKQVC 14

RESULT 5
ENTRY A57434 #type complete
TITLE protein-tyrosine kinase (EC 2.7.1.112) CAK-beta - rat

```



```

ALTERNATE_NAMES    cell adhesion kinase-beta
ORGANISM            #formal_name Rattus norvegicus #common_name Norway rat
DATE               08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change
19-Dec-1997
ACCESSIONS         A57434
REFERENCE          A57434
#authors           Sasaki, H.; Nagura, K.; Ishino, M.; Tobioaka, H.; Kotani, K.;
#journal           J. Biol. Chem. (1995) 270:21206-21219
#title             Cloning and characterization of cell adhesion kinase beta, a
                   novel protein-tyrosine kinase of the focal adhesion kinase
                   subfamily.
#cross-references  MUID:95403356
#accession         A57434
#status            preliminary
#molecule_type    mRNA
#residues          1-1009 #label SAS
#cross-references  GB:D45854; NID:g1000679; PID:d1008885; PID:g1000680
CLASSIFICATION     #superfamily unassigned Ser/Thr or Tyr-specific protein
                   kinases; protein kinase homology
KEYWORDS           ATP; phosphotransferase
FEATURE            423-686   #domain protein kinase homology #label KIN
                   431-439   #region protein kinase ATP-binding motif
SUMMARY            #length 1009 #molecular-weight 115725 #checksum 4981

Query Match      49.2%; Score 58; DB 2; Length 1009;
Best Local Similarity 42.9%; Pred. No. 1.56e+00;
Matches          6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 DVEKEDVRILKVCVF 45
QY 1 EINKDDVQHKQVCVF 14

RESULT 6
ENTRY 6
TITLE cell adhesion kinase beta - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change
17-Jul-1998
ACCESSIONS G02330
REFERENCE H01067
#authors   Sasaki, T.; Nagura, K.; Sasaki, H.
#submission submitted to the EMBL Data Library, December 1995
#accession G02330
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-1009 #label SAS
#cross-references EMBL:043522; NID:g1165218; PID:g1165219
GENETICS
#gene     CAKBeta
CLASSIFICATION #superfamily unassigned Ser/Thr or Tyr-specific protein
               kinases; protein kinase homology
FEATURE      423-686   #domain protein kinase homology #label KIN
               431-439   #region protein kinase ATP-binding motif
SUMMARY      #length 1009 #molecular-weight 115826 #checksum 4144

Query Match      49.2%; Score 58; DB 2; Length 1009;
Best Local Similarity 42.9%; Pred. No. 1.56e+00;
Matches          6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 DVEKEDVRILKVCVF 45
QY 1 EINKDDVQHKQVCVF 14

RESULT 7
ENTRY 7
TITLE tyrosine-specific protein kinase (EC 2.7.1.-) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 10-Apr-1996 #sequence_revision 27-Feb-1997 #text_change

```

```

17-Mar-1999
ACCESSIONS S60248
REFERENCE S60248
#authors   Lev, S.; Moreno, H.; Martinez, R.; Canoll, P.; Peles, E.;
           Musacchio, J.M.; Plozman, G.D.; Rudy, B.; Schlessinger, J.
#journal   Nature (1995) 376:737-745
#title     Protein tyrosine kinase PYK2 involved in Ca(2+)-induced
           regulation of ion channel and MAP kinase functions.
#cross-references MUID:95379967
#accession S60248
#status    preliminary; nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-1009 #label LEV
#cross-references EMBL:U33284; NID:g988304; PID:g988305
CLASSIFICATION #superfamily unassigned Ser/Thr or Tyr-specific protein
               kinases; protein kinase homology
KEYWORDS      phosphotransferase
FEATURE       423-686   #domain protein kinase homology #label KIN
               431-439   #region protein kinase ATP-binding motif
SUMMARY       #length 1009 #molecular-weight 115874 #checksum 3790

Query Match      49.2%; Score 58; DB 2; Length 1009;
Best Local Similarity 42.9%; Pred. No. 1.56e+00;
Matches          6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 DVEKEDVRILKVCVF 45
QY 1 EINKDDVQHKQVCVF 14

RESULT 8
ENTRY 8
TITLE gob-5 protein - Mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 18-Mar-1999 #sequence_revision 18-Mar-1999 #text_change
18-Mar-1999
ACCESSIONS JG0168
REFERENCE JG0168
#authors   Koniya, T.; Tanigawa, Y.; Hirohashi, S.
#journal   Biochem. Biophys. Res. Commun. (1999) 255:347-351
#title     Cloning and identification of the gene gob-5, which is
           expressed in intestinal goblet cells in mice.
#accession JG0168
#status    preliminary
#residues 1-913 #label KOM
#cross-references DBJ:AB016592
SUMMARY      #length 913 #molecular-weight 100070 #checksum 8755

Query Match      48.3%; Score 57; DB 3; Length 913;
Best Local Similarity 77.8%; Pred. No. 2.43e+00;
Matches          7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 720 VQDKQLCFS 728
QY 7 VQHKQVCFS 15

RESULT 9
ENTRY 9
TITLE alpha.alpha-trehalose-phosphate synthase (UDP-forming) (EC
           2.4.1.15) - Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change
20-Mar-1998
ACCESSIONS I83402
REFERENCE I83402; I60301
#authors   Kaasen, I.; McDougall, J.; Strom, A.R.
#journal   Gene (1994) 145:9-15
#title     Analysis of the otsBA operon for osmoregulatory trehalose
           synthesis in Escherichia coli and homology of the OtsA and
           OtsB proteins to the yeast trehalose-6-phosphate
           synthase/phosphatase complex.

```

```
##cross-references MUID:94320793
#accession I83402
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-474 #label RES
##cross-references EMBL:X69160; NID:g457774; PID:g603543
##experimental_source strain K-12; substrain CSH7
REFERENCE A64720
#authors
Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, M.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession H64952
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-474 #label BLAT
##cross-references GB:AF000283; GB:U00096; NID:g1788200; PID:g1788206;
UWGP:b1896
##experimental_source strain K-12, substrain MG1655
REFERENCE S33582
#authors McDougall, J.; Kaasen, I.; Strom, A.R.
#journal FEMS Microbiol. Lett. (1993) 107:25-30
#title A yeast gene for trehalose-6-phosphate synthase and its
complementation of an Escherichia coli otsA mutant.
#cross-references MUID:93224004
#accession S33584
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 'MT', 1-473 #label MCD
##cross-references the nucleotide sequence was submitted to the EMBL Data
Library
GENETICS
#gene otsA
#keywords glycosyltransferase; hexosyltransferase
SUMMARY #length 474 #molecular-weight 53611 #checksum 7682
Query Match 47.5%; Score 56; DB 2; Length 474;
Best Local Similarity 53.8%; Pred. No. 3.75e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 435 IVKNDINHWOECF 447
| | | | |
| | | | |
QY 2 INKDDVOHKQVCF 14
RESULT 10
ENTRY #type complete
TITLE phosphotransferase system enzyme II (EC 2.7.1.69),
glucose-specific, factor IIB - Staphylococcus carnosus
ALTERNATE_NAMES glucose permease; phosphoenolpyruvate:glucose
phosphotransferase system enzyme II, glucose-specific;
protein-Npi-phosphohistidine--sugar phosphotransferase,
glucose-specific, factor II
ORGANISM #formal_name Staphylococcus carnosus
DATE 06-Jan-1995 #sequence_revision 31-Jan-1997 #text_change
16-Dec-1998
ACCESSIONS S63606; S46953
REFERENCE S63605
#authors Christiansen, I.; Hengstenberg, W.
#journal Mol. Gen. Genet. (1996) 250:375-379
#title Cloning and sequencing of two genes from Staphylococcus
carnosus coding for glucose-specific PTS and their
expression in Escherichia coli K-12.
#cross-references MUID:96180655
#accession S63606
#status preliminary
#molecule_type DNA
##cross-references MUID:94320793
#accession I83402
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-474 #label RES
##cross-references EMBL:X69160; NID:g457774; PID:g603543
##experimental_source strain K-12; substrain CSH7
REFERENCE A64720
#authors
Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, M.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession H64952
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-474 #label BLAT
##cross-references GB:AF000283; GB:U00096; NID:g1788200; PID:g1788206;
UWGP:b1896
##experimental_source strain K-12, substrain MG1655
REFERENCE S33582
#authors McDougall, J.; Kaasen, I.; Strom, A.R.
#journal FEMS Microbiol. Lett. (1993) 107:25-30
#title A yeast gene for trehalose-6-phosphate synthase and its
complementation of an Escherichia coli otsA mutant.
#cross-references MUID:93224004
#accession S33584
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 'MT', 1-473 #label MCD
##cross-references the nucleotide sequence was submitted to the EMBL Data
Library
GENETICS
#gene otsA
#keywords glycosyltransferase; hexosyltransferase
SUMMARY #length 474 #molecular-weight 53611 #checksum 7682
Query Match 47.5%; Score 56; DB 2; Length 474;
Best Local Similarity 53.8%; Pred. No. 3.75e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 435 IVKNDINHWOECF 447
| | | | |
| | | | |
QY 2 INKDDVOHKQVCF 14
RESULT 10
ENTRY #type complete
TITLE phosphotransferase system enzyme II (EC 2.7.1.69),
glucose-specific, factor IIB - Staphylococcus carnosus
ALTERNATE_NAMES glucose permease; phosphoenolpyruvate:glucose
phosphotransferase system enzyme II, glucose-specific;
protein-Npi-phosphohistidine--sugar phosphotransferase,
glucose-specific, factor II
ORGANISM #formal_name Staphylococcus carnosus
DATE 06-Jan-1995 #sequence_revision 31-Jan-1997 #text_change
16-Dec-1998
ACCESSIONS S63606; S46953
REFERENCE S63605
#authors Christiansen, I.; Hengstenberg, W.
#journal Mol. Gen. Genet. (1996) 250:375-379
#title Cloning and sequencing of two genes from Staphylococcus
carnosus coding for glucose-specific PTS and their
expression in Escherichia coli K-12.
#cross-references MUID:96180655
#accession S63606
#status preliminary
#molecule_type DNA
##cross-references MUID:94320793
#accession I83402
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-474 #label RES
##cross-references EMBL:X69160; NID:g457774; PID:g603543
##experimental_source strain K-12; substrain CSH7
REFERENCE A64720
#authors
Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, M.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references MUID:97426617
#accession H64952
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-474 #label BLAT
##cross-references GB:AF000283; GB:U00096; NID:g1788200; PID:g1788206;
UWGP:b1896
##experimental_source strain K-12, substrain MG1655
REFERENCE S33582
#authors McDougall, J.; Kaasen, I.; Strom, A.R.
#journal FEMS Microbiol. Lett. (1993) 107:25-30
#title A yeast gene for trehalose-6-phosphate synthase and its
complementation of an Escherichia coli otsA mutant.
#cross-references MUID:93224004
#accession S33584
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 'MT', 1-473 #label MCD
##cross-references the nucleotide sequence was submitted to the EMBL Data
Library
GENETICS
#gene otsA
#keywords glycosyltransferase; hexosyltransferase
SUMMARY #length 474 #molecular-weight 53611 #checksum 7682
Query Match 47.5%; Score 56; DB 2; Length 474;
Best Local Similarity 53.8%; Pred. No. 3.75e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 435 IVKNDINHWOECF 447
| | | | |
| | | | |
QY 2 INKDDVOHKQVCF 14
RESULT 11
ENTRY #type complete
TITLE finger protein sdc-1 - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change
09-Sep-1997
ACCESSIONS S15093; A33165
REFERENCE S15093
#authors Nonet, M.L.; Meyer, B.J.
#journal Nature (1991) 351:65-68
#title Early aspects of Caenorhabditis elegans sex determination and
dosage compensation are regulated by a zinc-finger protein.
#cross-references MUID:91226537
#accession S15093
#status preliminary
#molecule_type mRNA
#residues 1-1203 #label NON
##cross-references EMBL:X58520; NID:g6848; PID:g6849
GENETICS
#gene sdc-1
#keywords DNA binding; nucleus; transcription regulation; zinc finger
SUMMARY #length 1203 #molecular-weight 139286 #checksum 6221
Query Match 47.5%; Score 56; DB 2; Length 1203;
Best Local Similarity 55.8%; Pred. No. 3.75e+00;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 938 RHREMCFSR 946
| | | | |
| | | | |
QY 8 QHKQVCFR 16
```

```

Best Local Similarity 50.0%; Pred.No. 5.78e+00;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 102 EEVQIGDVCFKR 113
    :||| :|||
QY 5 DDVQHKVCFSR 16

RESULT 14
ENTRY A46212 #type complete
TITLE MEK kinase - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
19-Dec-1997
ACCESSIONS A46212
REFERENCE A46212
#authors Lange-Carter, C.A.; Pleiman, C.M.; Gardner, A.M.; Blumer, K.J.; Johnson, G.L.
#journal Science (1993) 260:315-319
#title A divergence in the MAP kinase regulatory network defined by MEK kinase and Raf.
#cross-references MUID:93227040
#accession A46212
##status preliminary; not compared with conceptual translation
##molecule_type nucleic acid
##residues 1-687 #label LAN
##experimental_source brain
##note sequence extracted from NCBI backbone (NCBIP:129292)
CLASSIFICATION #superfamily unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
FEATURE
416-683 #domain protein kinase homology #label KIN\
424-432 #region protein kinase ATP-binding motif
SUMMARY #length 687 #molecular_weight 74702 #checksum 2264

Query Match 46.68; Score 55; DB 2; Length 687;
Best Local Similarity 53.8%; Pred.No. 5.78e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 340 DTYKDDVNHQKC 352
    : ||||| :|||
QY 1 EINKDDVQHKQVC 13

RESULT 15
ENTRY A54767 #type complete
TITLE transcription factor CPH1 - yeast (Candida albicans)
ALTERNATE_NAMES acid proteinase cross-reactive protein ACPR
ORGANISM #formal_name Candida albicans
DATE 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change
17-Mar-1999
ACCESSIONS A54767; A55246
REFERENCE A54767
#authors Maatathi, K.; Ganesan, K.; Datta, A.
#journal J. Biol. Chem. (1994) 269:22945-22951
#title Identification of a putative transcription factor in Candida albicans that can complement the mating defect of Saccharomyces cerevisiae stel2 mutants.
#cross-references MUID:94364983
#accession A54767
##molecule_type DNA
##residues 1-699 #label MAL
##cross-references GB:L16451
##note the authors suggest that the codon TAG at position 657 is not a terminator

REFERENCE A55246
#authors Liu, H.; Koehler, J.; Fink, G.R.
#journal Science (1994) 266:1723-1726
#title Suppression of hyphal formation in Candida albicans by mutation of a Stel2 homolog.
#accession A55246
##molecule_type DNA
##residues 1-481, 'M', 483-509, 'T', 511-609, 'N', 611-650 #label LIU

```

```
##cross-references GB:U15152
#note      nucleotide sequence not given; amino acid sequence not
           complete

GENETICS
#gene      CPH1; ACPR
KEYWORDS   DNA binding; transcription factor
SUMMARY    #length 699 #molecular-weight 78425 #checksum 8759

Query Match 46.6%; Score 55; DB 2; Length 699;
Best Local Similarity 40.0%; Pred. No. 5.78e+00;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 646 INKEEVKQSQTXMNR 660
Oy 2 INKDDVQHKQVCFSR 16
   ||::||:|:|
Search completed: Mon Aug 23 13:21:14 1999
Job time : 15 secs.
```

\*\*\*\*\*

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U33284; G988305; -  
 CC EMBL: L49207; G1082034; -  
 CC EMBL: U43522; G1165219; -  
 CC MIM: 601212; -  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PFAM: PF00069; pkinase; 1.  
 CC HSP: P11362; IFGI.  
 CC TYROSINE-PROTEIN KINASE; TRANSFERASE; ATP-BINDING; PHOSPHORYLATION.  
 CC DOMAIN 425 683 PROTEIN KINASE.  
 CC FT NP\_BIND 431 439 ATP (BY SIMILARITY).  
 CC FT BINDING 457 457 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 549 549 BY SIMILARITY.  
 CC FT MOD\_RES 579 579 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC FT CONFLICT 23 23 A -> G (IN REF. 3).  
 CC FT CONFLICT 256 256 G -> P (IN REF. 2).  
 CC FT CONFLICT 435 435 F -> L (IN REF. 3).  
 CC FT CONFLICT 780 780 R -> G (IN REF. 2).  
 CC SQ SEQUENCE 1009 AA; 115974 MW; 57140A7A CRC32;

Query Match 49.2%; Score 58; DB 1; Length 1009;

Best Local Similarity 42.9%; Pred. No. 4.40e-01; Indels 0; Gaps 0;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 DVEKEDVRLKVCVF 45

QY 1 EINKDDVQHKVCVF 14

RESULT 2

ID FAK2\_RAT STANDARD; PRT; 1009 AA.  
 AC P70600; Q63201;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE FOCAL ADHESION KINASE 2 (EC 2.7.1.112) (FAK2 2) (CELL ADHESION KINASE  
 DE BETA) (CAK BETA) (CALCIUM-DEPENDENT TYROSINE KINASE).  
 GN FAK2.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 97094711.  
 RA YU H., LI X., MARCHETTO G.S., DY R., HUNTER D., DAMSON T., WILM M.,  
 RA ANDEREGG R., GRAVES L., EARP H.S.;  
 RT "Activation of a novel calcium-dependent protein-tyrosine kinase.  
 RT Correlation with c-Jun N-terminal kinase but not mitogen-activated  
 RT protein kinase activation."  
 RL J. BIOL. CHEM. 271:29993-29998(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;  
 RX MEDLINE; 95403356.  
 RA SASAKI H., NAGURA K., ISHINO M., TOBIOKA H., KOTANI K., SASAKI T.;  
 RT "Cloning and characterization of cell adhesion kinase beta, a novel  
 RT protein-tyrosine kinase of the focal adhesion kinase subfamily."  
 RL J. BIOL. CHEM. 270:21206-21219(1995).  
 CC -!- FUNCTION: INVOLVED IN CALCIUM INDUCED REGULATION OF ION CHANNEL  
 CC AND ACTIVATION OF THE MAP KINASE SIGNALING PATHWAY. MAY REPRESENT  
 CC AN IMPORTANT SIGNALING INTERMEDIATE BETWEEN NEUROPEPTIDE ACTIVATED  
 CC RECEPTORS OR NEUROTRANSMITTERS THAT INCREASE CALCIUM FLUX AND THE  
 CC DOWNSTREAM SIGNALS THAT REGULATE NEURONAL ACTIVITY.  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +

CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
 CC -!- PTM: PHOSPHORYLATED ON TYROSINES.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: U69109; G1654377; -  
 CC EMBL: D45854; G1000680; -  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PFAM: PF00069; pkinase; 1.  
 CC HSP: P11362; IFGI.  
 CC TYROSINE-PROTEIN KINASE; TRANSFERASE; ATP-BINDING; PHOSPHORYLATION.  
 CC DOMAIN 425 683 PROTEIN KINASE.  
 CC FT NP\_BIND 431 439 ATP (BY SIMILARITY).  
 CC FT BINDING 457 457 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 549 549 BY SIMILARITY.  
 CC FT MOD\_RES 579 579 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC FT CONFLICT 205 205 E -> A (IN REF. 2).  
 CC SQ SEQUENCE 1009 AA; 115784 MW; 99C6DB44 CRC32;

Query Match 49.2%; Score 58; DB 1; Length 1009;

Best Local Similarity 42.9%; Pred. No. 4.40e-01; Indels 0; Gaps 0;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 32 DVEKEDVRLKVCVF 45

QY 1 EINKDDVQHKVCVF 14

RESULT 3

ID OTSA\_ECOLI STANDARD; PRT; 473 AA.  
 AC P31677;  
 DT 01-JUL-1993 (REL. 26, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE ALPHA, ALPHA-TREHALOSE-PHOSPHATE SYNTHASE [UDP-FORMING] (EC 2.4.1.15)  
 DE (TREHALOSE-6-PHOSPHATE SYNTHASE) (UDP-GLUCOSE-GLUCOSEPHOSPHATE  
 DE GLUCOSYLTRANSFERASE).  
 GN OTSA.  
 OS ESCHERICHIA COLI.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC ESCHERICHIA.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=K12 / CSH7;  
 RX MEDLINE; 94320793.  
 RA KAASEN I., MCDUGALL J., STROEM A.R.;  
 RT "Analysis of the otsBA operon for osmoregulatory trehalose synthesis  
 RT in Escherichia coli and homology of the OtsA and OtsB proteins to the  
 RT yeast trehalose-6-phosphate synthase/phosphatase complex."  
 RL GENE 145:9-15(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE; 97426617.  
 RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
 RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
 RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
 MAU B., SHAO Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL SCIENCE 277:1453-1474(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN-K12;  
 RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,  
 RA IKEMOTO K., INADA T., ISONO K., KIMURA S., ITOH T., KANAI K.,  
 RA KASAI H., KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M.,  
 RA KITAKAWA M., MAKINO K., MASUDA S., MIKI T., MIZOBUCHI K., MORI H.,  
 RA MOTOMURA K., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,  
 RA SAITO N., SAMPEI G., SEKI Y., TAGAMI H., TAKEMOTO K., WADA C.,  
 RA YAMAMOTO Y., YANO M.,  
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RP [14]  
 RP SEQUENCE OF 429-473 FROM N.A.  
 RC STRAIN-K12 / EMG2;  
 RA ESTEP P., O'KEEFE T., ROBISON K., CHURCH G.M.,  
 RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- CATALYTIC ACTIVITY: UDP-GLUCOSE + D-GLUCOSE 6-PHOSPHATE = UDP +  
 CC ALPHA,ALPHA-TREHALOSE 6-PHOSPHATE.  
 CC -!- SIMILARITY: TO THE YEAST ENZYME (GENE TPS1 OR C1F1).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X69160; G603543; -;  
 DR EMBL: AF000283; G1788206; -;  
 DR EMBL: D90831; G1736555; ALT\_INIT.  
 DR EMBL: U27211; G862973; -;  
 DR PIR: S33584; S33584.  
 DR ECGENE: EG11751; OTSA.  
 DR PFAM: PF00982; Trehalosep\_syn; 1.  
 KW TRANSFERASE; GLYCOSYLTRANSFERASE.  
 FT INIT\_MET 0  
 FT SEQUENCE 473 AA; 53480 MW; 5C146A05 CRC32;

Query Match 47.58; Score 56; DB 1; Length 473;  
 Best Local Similarity 53.8%; Pred. No. 1.18e+00;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 434 IVKNDINHWOECF 446  
 | | | | | | | | | |  
 QY 2 INKDDVQHKVCF 14

RESULT 4  
 ID SDC1 CAPEL STANDARD; PRT; 1203 AA.  
 AC P24349; Q20672;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE ZINC FINGER PROTEIN SDC-1.  
 GN SDC-1 OR EGL-16 OR F52E10.1.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTERA; RHABDITIDA; RHABDITIDA;  
 OC RHABDITIDA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE: 91226537.  
 RA NONET M.L., MEYER B.J.;  
 RT "Early aspects of Caenorhabditis elegans sex determination and dosage  
 RT compensation are regulated by a zinc-finger protein.";  
 RL NATURE 351:65-68(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA LLOYD C.;  
 RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- FUNCTION: EMBRYONIC TRANSCRIPTION FACTOR REGULATING DOWNSTREAM  
 CC GENES INVOLVED SPECIFICALLY IN THE SEX DETERMINATION AND DOSAGE  
 CC COMPENSATION PATHWAYS, OR REGULATING OTHER GENES INVOLVED IN THE

CC COORDINATE CONTROL OF BOTH PROCESSES.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: X58520; G6849; -;  
 DR EMBL: Z54282; E1346984; -;  
 DR PIR: A33165; A33165.  
 DR PIR: S15093; S15093.  
 DR WORPEP: F52E10.1; CE03393.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 2.  
 DR PFAM: PF00096; zf-C2H2; 4.  
 DR TRANSFAC: T00742; -;  
 KW SEXUAL DIFFERENTIATION; TRANSCRIPTION REGULATION; ZINC-FINGER;  
 KW METAL-BINDING; DNA-BINDING; DEVELOPMENTAL PROTEIN; NUCLEAR PROTEIN.  
 FT DOMAIN 3 15  
 FT ZN\_FING 117 139 ASP/GLU-RICH (ACIDIC).  
 FT ZN\_FING 145 168 C2H2-TYPE.  
 FT ZN\_FING 235 256 C2H2-TYPE.  
 FT ZN\_FING 270 292 C2H2-TYPE.  
 FT ZN\_FING 488 515 C2H2-TYPE.  
 FT ZN\_FING 523 545 C2H2-TYPE.  
 FT ZN\_FING 654 676 C2H2-TYPE.  
 FT CONFLICT 16 16 V -> D (IN REF. 2).  
 FT CONFLICT 944 945 FS -> KT (IN REF. 2).  
 SQ SEQUENCE 1203 AA; 139286 MW; 8A760914 CRC32;

Query Match 47.5%; Score 56; DB 1; Length 1203;  
 Best Local Similarity 55.6%; Pred. No. 1.18e+00;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 938 RHREMCFSR 946  
 | | | | | | | | | |  
 QY 8 QHKQVCFSR 16

RESULT 5  
 ID AFUC\_HAEIN STANDARD; PRT; 328 AA.  
 AC P44531;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PUTATIVE FERRIC TRANSPORT ATP-BINDING PROTEIN AFUC.  
 GN AFUC OR H10126.  
 OS HAEMOPHILUS INFLUENZAE.  
 OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;  
 OC HAEMOPHILUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20;  
 RX MEDLINE: 95350630.  
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
 RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
 RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
 RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,  
 RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL SCIENCE 269:496-512(1995).  
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR IRON; PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE  
 CC TRANSPORT SYSTEM.  
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY

```
CC -----
CC (ABC TRANSPORTERS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32698; G1573081; -
DR TIGR; H10126; -
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
DR IRON_TRANSPORT; TRANSPORT; ATP-BINDING.
FT NP_BIND 39 46 ATP (POTENTIAL).
SQ SEQUENCE 328 AA; 36046 MW; 6F25D726 CRC32;

Query Match 46.6%; Score 55; DB 1; Length 328;
Best Local Similarity 21.4%; Pred. No. 1.91e+00;
Matches 3; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Db 69 DVTKSSIQNRDICI 82
QY 1 EINKDDVQHKQVCF 14

RESULT 6
ID BCHX_RHOSH STANDARD; PRT; 333 AA.
AC Q02431;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CHLOROPHYLLIDE REDUCTASE 35.5 KD CHAIN (EC 1.3.1.1-) (CHLORIN
DE REDUCTASE).
GN BCHX.
OS RHODOBACTER SPHAEROIDES (RHODOPSEUDOMONAS SPHAEROIDES).
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHODOBACTER GROUP;
OC RHODOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCIB 8253.
RX MEDLINE; 93173096.
RA MCGLYNN P., HUNTER C.N.;
RT "Genetic analysis of the bchC and bchA genes of Rhodobacter
RT sphaeroides."
RL MOL. GEN. GENET. 236:227-234(1993).
CC -!- FUNCTION: CONVERTS CHLOROPHYLLS (CHL) INTO BACTERIOCHLOROPHYLLS
CC (BCHL) BY REDUCING RING B OF THE TETRAPYRROLE.
CC -!- PATHWAY: BACTERIOCHLOROPHYLL BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE NIFH/FRXC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68795; G49172; -
DR PIR; S28270; S28270.
DR PIR; S30915; S30915.
DR PROSITE; PS00692; NIFH_FRXC_2; 1.
DR PROSITE; PS00746; NIFH_FRXC_1; 1.
DR PFAM; PF00142; Iser4_Nifh; 1.
DR HSP; P00459; IN2C.
KW OXIDOREDUCTASE; PHOTOSYNTHESIS; CHLOROPHYLL BIOSYNTHESIS; ATP-BINDING;
KW IRON-SULFUR; 4FE-4S.
FT NP_BIND 42 49 ATP (BY SIMILARITY).
FT METAL 130 130 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 165 165 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 333 AA; 35575 MW; 252CCB5F CRC32;
```

```
Query Match 46.6%; Score 55; DB 1; Length 333;
Best Local Similarity 50.0%; Pred. No. 1.91e+00;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 102 EEVQIGDVCFKR 113
QY 5 DDVQHKQVCFSR 16

RESULT 7
ID AFUC_ACTPL STANDARD; PRT; 348 AA.
AC Q57293;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE FERRIC TRANSPORT ATP-BINDING PROTEIN AFUC.
GN AFUC.
OS ACTINOBACILLUS PLEUROPNEUMONIAE (HAEMOPHILUS PLEUROPNEUMONIAE).
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC ACTINOBACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S 4074 / SEROTYPE 1, AND K17 / SEROTYPE 5;
RX MEDLINE; 96401417.
RA CHIN N., FREY J., CHANG C.F., CHANG Y.F.;
RT "Identification of a locus involved in the utilization of iron by
RT Actinobacillus pleuropneumoniae."
RL FEMS MICROBIOL. LETT. 143:1-6(1996).
CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR IRON; PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
CC TRANSPORT SYSTEM.
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U04954; G1477453; -
DR EMBL; U05042; G1469284; -
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR PFAM; PF00005; ABC_tran; 1.
KW IRON_TRANSPORT; TRANSPORT; ATP-BINDING.
FT NP_BIND 38 45
SQ SEQUENCE 348 AA; 38691 MW; 698CE1EE CRC32;

Query Match 46.6%; Score 55; DB 1; Length 348;
Best Local Similarity 21.4%; Pred. No. 1.91e+00;
Matches 3; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Db 68 DVTKSSIQNRDICI 81
QY 1 EINKDDVQHKQVCF 14

RESULT 8
ID MEK1_MOUSE STANDARD; PRT; 834 AA.
AC P53349; G060831;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MAPK/ERK KINASE KINASE 1 (EC 2.7.1.1-) (MEK KINASE 1) (MEKK 1)
DE (FRAGMENT).
GN MEKK1 OR MEKK.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MORIDAE; MURINAE; MUS.
RN [1]
```



```
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-BRAIN;
RX MEDLINE: 93227040.
RA LANGE-CARTER C.A., PLEIMAN C.M., GARDNER A.M., BLUMER K.J.,
RT JOHNSON G.L.;
RA "A divergence in the MAP kinase regulatory network defined by MEK
RT kinase and Raf.";
RL SCIENCE 260:315-319(1993).
RN [2]
RP SEQUENCE OF 137-834 FROM N.A.
RC STRAIN-BALB/C; TISSUE-HEART;
RA WHITMARSH A.J., SHORE P., SHARROCKS A.D., DAVIS R.J.;
RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE MAPKK 1 AND MAPKK 2
CC (MEK1/MEK2) WHICH LEADS TO PHOSPHORYLATION OF MAP KINASES.
CC POTENTIALLY ACTIVATES THE JNK/SAPK GROUP OF MAP KINASES.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE HEART AND SPLEEN WHILE
CC A LOWER LEVEL EXPRESSION IS SEEN IN THE LIVER.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE MAP KINASE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13103; G293730; ALT_INIT.
DR EMBL: U23470; G1143820; -.
DR MGD; MGI:104845; MEKK.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR HSSP; P24941; IAQI.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW PHOSPHORYLATION.
FT NON_TER 1 1
FT DOMAIN 565 830 PROTEIN KINASE.
FT NP_BIND 571 578 ATP (BY SIMILARITY).
FT BINDING 594 594 ATP (BY SIMILARITY).
FT ACT_SITE 691 691 BY SIMILARITY.
FT CONFLICT 224 224 L -> V (IN REF. 1).
FT CONFLICT 808 808 L -> V (IN REF. 1).
SQ SEQUENCE 834 AA; 90878 MW; D48C9CC4 CRC32;
Query Match 46.6%; Score 55; DB 1; Length 834;
Best Local Similarity 53.8%; Pred. No. 1.91e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 487 DTYKDDVNHQKC 499
QY : |||||:|:|
1 EINKDDVQHKQVC 13
RESULT 9
ID MEK1_RAT STANDARD; PRT; 1493 AA.
AC Q82925;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MAPK/ERK KINASE KINASE 1 (EC 2.7.1.-) (MEK KINASE 1) (MEKK 1).
GN MEKK1 OR MEKK.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RC TISSUE-BRAIN;
RX MEDLINE: 96224276.
```

```
RA XU S., ROBBINS D.J., CHRISTERSON L.B., ENGLISH J.M.,
RA VANDERBILT C.A., COBB M.H.;
RT "Cloning of rat MEK kinase 1 cDNA reveals an endogenous membrane-
associated 195-kDa protein with a large regulatory domain.";
RL PROC. NATL. ACAD. SCI. U.S.A. 93:5291-5295(1996).
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE MAPKK 1 AND MAPKK 2
CC (MEK1/MEK2) WHICH LEADS TO PHOSPHORYLATION OF MAP KINASES. MOST
CC POTENTIALLY ACTIVATES THE JNK/SAPK GROUP OF MAP KINASES, AND LESS
CC EFFICIENTLY ERK2 OR P38.
CC -!- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED.
CC -!- TISSUE SPECIFICITY: MOST HIGHLY EXPRESSED IN SPLEEN, KIDNEY, AND
CC LUNG.
CC -!- PTM: AUTOPHOSPHORYLATED.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE MAP KINASE KINASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 PH DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U48596; G1354137; -.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR HSSP; P24941; IAQI.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW PHOSPHORYLATION.
FT DOMAIN 2 5 POLY-ALA.
FT DOMAIN 17 20 POLY-ALA.
FT DOMAIN 25 29 POLY-GLY.
FT DOMAIN 74 149 PRO-RICH.
FT DOMAIN 111 114 POLY-PRO.
FT DOMAIN 233 291 PRO-RICH.
FT DOMAIN 412 421 POLY-SER.
FT DOMAIN 439 555 PH.
FT DOMAIN 643 750 PH.
FT DOMAIN 1163 1168 POLY-GLU.
FT DOMAIN 1197 1200 POLY-ILE.
FT DOMAIN 1224 1489 PROTEIN KINASE.
FT NP_BIND 1230 1237 ATP (BY SIMILARITY).
FT BINDING 1253 1253 ATP (BY SIMILARITY).
FT ACT_SITE 1350 1350 BY SIMILARITY.
FT MUTAGEN 1369 1369 D->A; INACTIVATION.
SQ SEQUENCE 1493 AA; 161315 MW; 982785D9 CRC32;
Query Match 46.6%; Score 55; DB 1; Length 1493;
Best Local Similarity 53.8%; Pred. No. 1.91e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 1146 DTYKDDVNHQKC 1158
QY : |||||:|:|
1 EINKDDVQHKQVC 13
RESULT 10
ID ARF4_MOUSE STANDARD; PRT; 179 AA.
AC P36403;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ADP-RIBOSYLATION FACTOR 4.
GN ARF4.
OS MUS MUSCULUS (MOUSE), AND RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC SPECIES-RAT; TISSUE=TESTIS;
RX MEDLINE; 94342342.
RA LEE F.J.S., STEVENS L.A., HALL L.M., MORTAGH J.J. JR., KAO Y.L.,
RR MISHIMA K., PRICE S.R., NIGHTINGALE M.S., KOUSVELARI E., MOSS J.,
RT VAUGHAN M.;
RT "Characterization of class II and class III ADP-ribosylation factor
RL genes and proteins in Drosophila melanogaster.";
RL J. BIOL. CHEM. 269:21555-21560(1994).
CC -!- FUNCTION: GTP-BINDING PROTEIN THAT FUNCTIONS AS AN ALLOSTERIC
CC RIBOSYLTRANSFERASE. INVOLVED IN PROTEIN TRAFFICKING; MAY MODULATE
CC VESICLE BUDDING AND UNCOATING WITHIN THE GOLGI APPARATUS.
CC -!- SIMILARITY: BELONGS TO THE ARF FAMILY OF GTP-BINDING PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L25062; G507232; -.
DR FLXBASE; FBgn0013749; Arf102F.
DR PROSITE; PS01019; ARF; 1.
DR PFAM; PF00025; arf; 1.
DR HSP; P32889; IRRF.
KW GTP-BINDING; MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;
KW GOLGI STACK.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (POTENTIAL).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
SQ SEQUENCE 179 AA; 20485 MW; 1973CECC CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;
Best Local Similarity 35.7%; Pred. No. 3.09e+00;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0.

Db 50 FNVTVEYKNCIFT 63
QY 2 INKDQVHKQCFS 15
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~

RESULT 12
ID ARF5_CHICK STANDARD; PRT; 179 AA.
AC P49702;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE ADP-RIBOSYLATION FACTOR 5.
GN ARF5 OR CPS3.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGRON; TISSUE=EMBRYO;
RA KONKEL D.A., SONG S.K.;
RL SUBMITTED (OCT-1990) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: GTP-BINDING PROTEIN THAT FUNCTIONS AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT, AN ADP-
CC RIBOSYLTRANSFERASE. INVOLVED IN PROTEIN TRAFFICKING; MAY MODULATE
CC VESICLE BUDDING AND UNCOATING WITHIN THE GOLGI APPARATUS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE ARF FAMILY OF GTP-BINDING PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L25062; G507232; -.
DR FLXBASE; FBgn0013749; Arf102F.
DR PROSITE; PS01019; ARF; 1.
DR PFAM; PF00025; arf; 1.
DR HSP; P32889; IRRF.
KW GTP-BINDING; MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;
KW GOLGI STACK.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (POTENTIAL).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
SQ SEQUENCE 179 AA; 20485 MW; 1973CECC CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;
Best Local Similarity 35.7%; Pred. No. 3.09e+00;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0.

Db 50 FNVTVEYKNCIFT 63
QY 2 INKDQVHKQCFS 15
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~

RESULT 11
ID ARF2_DROME STANDARD; PRT; 179 AA.
AC P40945;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ADP-RIBOSYLATION FACTOR 2.
GN ARF102F OR ARF2.
OS DROSOPHILA MELAN
```

```
-----
CC EMBL; X55998; G899264; -
DR PROSITE: PS01019; ARF; 1.
DR PFAM; PF00025; arf; 1.
DR HSP; P32889; IRRF.
KW GTP-BINDING: MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;
KW GOLGI STACK.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (POTENTIAL).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
SQ SEQUENCE 179 AA; 20348 MW; C173D690 CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;
Best Local Similarity 35.7%; Pred. No. 3.09e+00;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 50 FNVETVEYKNICFT 63
QY 2 INKDDVQHKVCFS 15

-----
RESULT 13
ID ARF4_XENLA STANDARD; PRT; 179 AA.
AC P51844;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE ADP-RIBOSYLATION FACTOR 4.
GN ARF4.
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPODEA; PIPODEA; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RA ROMAN A. L.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: GTP-BINDING PROTEIN THAT FUNCTIONS AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT, AN ADP-
CC RIBOSYLTRANSFERASE. INVOLVED IN PROTEIN TRAFFICKING; MAY MODULATE
CC VESICLE BUDDING AND UNCOATING WITHIN THE GOLGI APPARATUS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE ARF FAMILY OF GTP-BINDING PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U31460; G967251; -
DR PROSITE: PS01019; ARF; 1.
DR PFAM; PF00025; arf; 1.
DR HSP; P32889; IRRF.
KW GTP-BINDING: MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;
KW GOLGI STACK.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (POTENTIAL).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
SQ SEQUENCE 179 AA; 20454 MW; 71048EC3 CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;
Best Local Similarity 35.7%; Pred. No. 3.09e+00;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 50 FNVETVEYKNICFT 63
QY 2 INKDDVQHKVCFS 15
-----
```

```
-----
RESULT 14
ID ARF5_HUMAN STANDARD; PRT; 179 AA.
AC P26437;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ADP-RIBOSYLATION FACTOR 5.
GN ARF5.
OS HOMO SAPIENS (HUMAN); MUS MUSCULUS (MOUSE), AND
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=HUMAN;
RX MEDLINE; 91131565.
RA TSUCHIYA M., PRICE S.R., TSAI S.-C., MOSS J., VAUGHAN M.;
RT "Molecular identification of ADP-ribosylation factor mRNAs and their
RT expression in mammalian cells.";
RL J. BIOL. CHEM. 266:2772-2777(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=HUMAN;
RX MEDLINE; 97312710.
RA MCGUIRE R.E., DAIGER S.P., GREEN E.D.;
RT "Localization and characterization of the human ADP-ribosylation
RT factor 5 (ARF5) gene.";
RL GENOMICS 41:481-484(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=RAT; TISSUE=BRAIN;
RA NIGHTINGALE M.S., PRICE S.R., TSUCHIYA M., MOSS J., VAUGHAN M.;
RL SUBMITTED (XXX-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=MOUSE; STRAIN=ICR; TISSUE=BRAIN;
RX MEDLINE; 97103475.
RA HOSAKA M., TODA K., TAKATSU H., TORII S., MURAKAMI K., NAKAYAMA K.;
RT "Structure and intracellular localization of mouse ADP-ribosylation
RT factors type 1 to type 6 (ARF1-ARF6).";
RL J. BIOCHEM. 120:813-819(1996).
CC -!- FUNCTION: GTP-BINDING PROTEIN THAT FUNCTIONS AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT, AN ADP-
CC RIBOSYLTRANSFERASE. INVOLVED IN PROTEIN TRAFFICKING; MAY MODULATE
CC VESICLE BUDDING AND UNCOATING WITHIN THE GOLGI APPARATUS.
CC -!- SIMILARITY: BELONGS TO THE ARF FAMILY OF GTP-BINDING PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57567; G178987; -
DR EMBL; U73002; G2088529; -
DR EMBL; L12384; G438870; -
DR EMBL; D87902; G1565215; -
DR PIR; A23741; A23741.
DR MIM; 103188; -
DR MGD; MGI:99434; ARF5.
DR PROSITE: PS01019; ARF; 1.
DR PFAM; PF00025; arf; 1.
DR HSP; P32889; IRRF.
KW GTP-BINDING: MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;
KW GOLGI STACK.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (POTENTIAL).
FT NP_BIND 23 30 GTP (BY SIMILARITY).
FT NP_BIND 66 70 GTP (BY SIMILARITY).
```

FT NP\_BIND 125 128 GTP (BY SIMILARITY).  
SQ SEQUENCE 179 AA; 20398 MW; 5827490B CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;  
Best Local Similarity 35.7%; Pred. No. 3.09e+00;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 50 FNVTVEYKNICFT 63  
QY 2 INKDDVQHKQVCF 15

Db 50 FNVTVEYKNICFT 63  
QY 2 INKDDVQHKQVCF 15

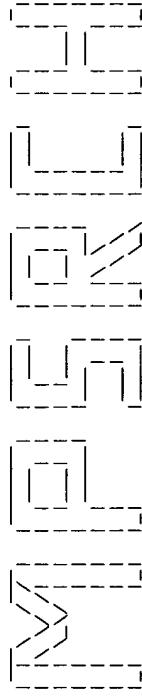
Search completed: Mon Aug 23 13:20:02 1999  
Job time : 10 secs.

## RESULT 15

ID ARF4\_HUMAN STANDARD; PRT; 179 AA.  
AC P18085; P21371;  
DT 01-MAY-1991 (REL. 18; CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE ADP-RIBOSYLATION FACTOR 4.  
GN ARF4.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90192776.  
RA MONACO L., MURTAGH J.J. JR., NEWMAN K.B., TSAI S.-C., MOSS J.,  
RA VAUGHAN M.;  
RT "Selective amplification of an mRNA and related pseudogene for a  
human ADP-ribosylation factor, a guanine nucleotide-dependent protein  
activator of cholera toxin.";  
RT activator of cholera toxin.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 87:2206-2210(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91115891.  
RA KAHN R.A., KERN F.G., CLARK J., GELMANN E.P., RULKA C.;  
RT "Human ADP-ribosylation factors. A functionally conserved family of  
GTP-binding proteins.";  
RL J. BIOL. CHEM. 266:2606-2614(1991).  
CC -!- FUNCTION: GTP-BINDING PROTEIN THAT FUNCTIONS AS AN ALLOSTERIC  
ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT. AN ADP-  
RIBOSYLTRANSFERASE. INVOLVED IN PROTEIN TRAFFICKING; MAY MODULATE  
VESICLE BUDDING AND UNCOATING WITHIN THE GOLGI APPARATUS.  
CC -!- SIMILARITY: BELONGS TO THE ARF FAMILY OF GTP-BINDING PROTEINS.  
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE ARF2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M36341; G178985; -  
DR PIR; A35091; A35091.  
DR PIR; B38622; B38622.  
DR MIN; 601177; -  
DR PROSITE; PS01019; ARF; 1.  
DR PFAM; PF00025; arf; 1.  
DR HSP; P32889; LHUR.  
KW GTP-BINDING; MULTIGENE FAMILY; MYRISTYLATION; PROTEIN TRANSPORT;  
KW GOLGI STACK.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT LIPID 1 1 MYRISTATE (POTENTIAL).  
FT NP\_BIND 23 30 GTP (BY SIMILARITY).  
FT NP\_BIND 66 70 GTP (BY SIMILARITY).  
FT NP\_BIND 125 128 GTP (BY SIMILARITY).  
SQ SEQUENCE 179 AA; 20379 MW; DC198A21 CRC32;

Query Match 45.8%; Score 54; DB 1; Length 179;  
Best Local Similarity 35.7%; Pred. No. 3.09e+00;  
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:20:20 1999; MasPar time 13.15 Seconds  
66.389 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-49  
Description: (1-16) from US09049596.pep  
Perfect Score: 118  
Sequence: 1 EINKDDVQHKQVCFSR 16

Scoring table: PAM 150  
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sp\_trembl  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 25.228; Variance 31.224; scale 0.808

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	66	55.9	913	11	088826	1.50e-02
2	61	51.7	367	2	087084	1.93e-01
3	61	51.7	768	11	035266	1.93e-01
4	61	51.7	831	2	024906	1.93e-01
5	60	50.8	251	2	034983	3.18e-01
6	59	50.0	309	2	060036	5.20e-01
7	58	49.2	150	4	014290	8.47e-01
8	58	49.2	317	5	044889	8.47e-01
9	58	49.2	501	5	044865	8.47e-01
10	56	47.5	631	5	044760	2.21e+00
11	56	47.5	692	2	033922	2.21e+00
12	55	46.6	136	2	052118	3.54e+00
13	55	46.6	347	5	018068	3.54e+00
14	55	46.6	467	10	081143	3.54e+00
15	55	46.6	473	5	016252	3.54e+00
16	55	46.6	604	4	043309	3.54e+00
17	55	46.6	1207	4	043724	3.54e+00
18	54	45.8	278	5	025679	5.63e+00
19	54	45.8	714	5	062436	5.63e+00
20	54	45.8	1310	5	024852	5.63e+00

21	54	45.8	2743	4	015162	POLYPOSIS LOCUS-ENCODE	5.63e+00
22	54	45.8 <td>2829</td> <td>13</td> <td>P70039</td> <td>ADENOMATOUS POLYPOSIS<th>5.63e+00</th></td>	2829	13	P70039	ADENOMATOUS POLYPOSIS <th>5.63e+00</th>	5.63e+00
23	54	45.8 <td>2842</td> <td>11</td> <td>P70478</td> <td>APC PROTEIN<th>5.63e+00</th></td>	2842	11	P70478	APC PROTEIN <th>5.63e+00</th>	5.63e+00
24	54	45.8 <td>2844</td> <td>4</td> <td>Q15163</td> <td>POLYPOSIS LOCUS-ENCODE<th>5.63e+00</th></td>	2844	4	Q15163	POLYPOSIS LOCUS-ENCODE <th>5.63e+00</th>	5.63e+00
25	54	45.8 <td>2845</td> <td>11</td> <td>Q61315</td> <td>ADENOMATOUS POLYPOSIS<th>5.63e+00</th></td>	2845	11	Q61315	ADENOMATOUS POLYPOSIS <th>5.63e+00</th>	5.63e+00
26	53	44.9 <td>26</td> <td>10</td> <td>Q38753</td> <td>ASP8 MADS BOX (FRAGMENT)<th>8.91e+00</th></td>	26	10	Q38753	ASP8 MADS BOX (FRAGMENT) <th>8.91e+00</th>	8.91e+00
27	53	44.9 <td>268</td> <td>2</td> <td>087805</td> <td>ELONGATION FACTOR EF-T<th>8.91e+00</th></td>	268	2	087805	ELONGATION FACTOR EF-T <th>8.91e+00</th>	8.91e+00
28	53	44.9 <td>280</td> <td>2</td> <td>084536</td> <td>ACYL-CARRIER UDP-GLCNA<th>8.91e+00</th></td>	280	2	084536	ACYL-CARRIER UDP-GLCNA <th>8.91e+00</th>	8.91e+00
29	53	44.9 <td>317</td> <td>10</td> <td>Q22050</td> <td>1-AMINOCYCLOPROPANE-1-<th>8.91e+00</th></td>	317	10	Q22050	1-AMINOCYCLOPROPANE-1- <th>8.91e+00</th>	8.91e+00
30	53	44.9 <td>325</td> <td>2</td> <td>Q51301</td> <td>STRAIN STOLF (HOST CAT<th>8.91e+00</th></td>	325	2	Q51301	STRAIN STOLF (HOST CAT <th>8.91e+00</th>	8.91e+00
31	53	44.9 <td>334</td> <td>2</td> <td>Q51299</td> <td>STRAIN AAY (HOST CATHA<th>8.91e+00</th></td>	334	2	Q51299	STRAIN AAY (HOST CATHA <th>8.91e+00</th>	8.91e+00
32	53	44.9 <td>335</td> <td>2</td> <td>Q51300</td> <td>STRAIN KV (HOST CATHAR<th>8.91e+00</th></td>	335	2	Q51300	STRAIN KV (HOST CATHAR <th>8.91e+00</th>	8.91e+00
33	53	44.9 <td>354</td> <td>2</td> <td>084846</td> <td>HYPOTHETICAL 40.1 KD P<th>8.91e+00</th></td>	354	2	084846	HYPOTHETICAL 40.1 KD P <th>8.91e+00</th>	8.91e+00
34	53	44.9 <td>712</td> <td>4</td> <td>Q75853</td> <td>WUGSC-H DJ0808A01.2 PR<th>8.91e+00</th></td>	712	4	Q75853	WUGSC-H DJ0808A01.2 PR <th>8.91e+00</th>	8.91e+00
35	53	44.9	1078	14	Q90193	RNA-DEPENDENT RNA POLY <th>8.91e+00</th>	8.91e+00
36	53	44.9	1229	10	Q49749	P-GLYCOPROTEIN-LIKE PR <th>8.91e+00</th>	8.91e+00
37	53	44.9	1993	5	P90670	SODIUM CHANNEL ALPHA-S <th>8.91e+00</th>	8.91e+00
38	52	44.1	26	10	Q38747	ASP13 MADS BOX (FRAGME <th>1.40e+01</th>	1.40e+01
39	52	44.1	231	1	P95952	ORF C04013	1.40e+01
40	52	44.1	259	5	Q24782	CNOX5 HOMEODOMAIN PROT <th>1.40e+01</th>	1.40e+01
41	52	44.1	328	5	Q45313	C54E10.2 PROTEIN. <th>1.40e+01</th>	1.40e+01
42	52	44.1	388	5	Q94739	CYCLIN B. <th>1.40e+01</th>	1.40e+01
43	52	44.1	417	5	Q94740	CYCLIN B. <th>1.40e+01</th>	1.40e+01
44	52	44.1	695	11	Q35272	PUTATIVE PHEROMONE REC <th>1.40e+01</th>	1.40e+01
45	52	44.1	802	11	Q35271	PUTATIVE PHEROMONE REC <th>1.40e+01</th>	1.40e+01

## ALIGNMENTS

RESULT 1  
ID 088826 PRELIMINARY; PRT; 913 AA.  
AC 088826;  
DT 01-NOV-1998 (TREMREL. 08, CREATED)  
DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE GOB-5 PROTEIN.  
GN GOB-5.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RA KOMIYA T., TANIGAWA Y., HIROHASHI S.;  
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet  
cells in mice."  
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AB017156; D1034712;  
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 55.9%; Score 66; DB 11; Length 913;  
Best Local Similarity 80.0%; Pred. No. 1.50e-02;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 720 VODKQCFSR 729

QY 7 VQHKQVCFSR 16

RESULT 2

ID 087084 PRELIMINARY; PRT; 367 AA.  
AC 087084;  
DT 01-NOV-1998 (TREMREL. 08, CREATED)  
DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE CELL DIVISION-RELATED GENE.  
GN CDRA.  
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).  
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;  
OC HELICOBACTER.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPK5;

RX MEDLINE: 98422487.  
 RA TAKEUCHI H., SHIRAI M., AKADA J.K., TSUDA M., NAKAZAWA T.;  
 RT "Nucleotide sequence and characterization of *cdra*, a cell  
 RL J. BACTERIOL. 180:5263-5268(1998).  
 DR EMBL: AB003309; D1034465; -.  
 KW CELL DIVISION.  
 SQ SEQUENCE 367 AA; 42500 MW; 7969DA1B CRC32;

Query Match 51.7%; Score 61; DB 2; Length 367;  
 Best Local Similarity 66.7%; Pred. No. 1.93e-01;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 164 DINHKEVCF 172  
 I: |||||  
 QY 6 DVQHKQVCF 14

RESULT 3  
 ID O35266 PRELIMINARY; PRT: 768 AA.  
 AC O35266;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE PUTATIVE PHEROMONE RECEPTOR.  
 GN GO-VN2.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
 OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE: 97433086.  
 RA HERRADA G., DULAC C.;  
 RT "A novel family of putative pheromone receptors in mammals with a  
 RT topographically organized and sexually dimorphic distribution.";  
 RL CELL 90:763-773(1997).  
 DR EMBL: AF016179; G2367607; -.  
 DR PFAM: PF00003; 7tm3; 1.  
 SQ SEQUENCE 768 AA; 88275 MW; 5E1AF896 CRC32;

Query Match 51.7%; Score 61; DB 11; Length 768;  
 Best Local Similarity 41.7%; Pred. No. 1.93e-01;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 145 KKETONKEICFA 156  
 I: |||||  
 QY 4 KDVQHKQVCF 15

RESULT 4  
 ID O24906 PRELIMINARY; PRT: 831 AA.  
 AC O24906;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE CONSERVED HYPOTHETICAL ATP-BINDING PROTEIN.  
 GN HP0066.  
 OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).  
 OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;  
 OC HELICOBACTER.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695;  
 RX MEDLINE: 97394467.  
 RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
 RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
 RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
 RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
 RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,  
 RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
 RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATHEY L., WALLIN E.,  
 RA HAYES W.S., BORODOVSKIY M., KARP P.D., SMITH H.O., FRASER C.M.,

RA VENTER J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
 RT *pylori* [published erratum appears in Nature 1997 Sep  
 RT 25:389(6649):412].";  
 RL NATURE 388:539-547(1997).  
 DR EMBL: AE000528; G2313138; -.  
 DR TIGR: HP0066; -.  
 KW HYPOTHETICAL PROTEIN; ATP-BINDING.  
 SQ SEQUENCE 831 AA; 95646 MW; 5CE1347C CRC32;

Query Match 51.7%; Score 61; DB 2; Length 831;  
 Best Local Similarity 66.7%; Pred. No. 1.93e-01;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 352 DINHKEVCF 360  
 I: |||||  
 QY 6 DVQHKQVCF 14

RESULT 5  
 ID O34983 PRELIMINARY; PRT: 251 AA.  
 AC O34983;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE RAS-LIKE PROTEIN HOMOLOG.  
 GN YOAP.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH D.;  
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE: 98044033.  
 RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
 RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,  
 RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
 RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
 RA CHOI S.-K., CODANI J.-J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
 RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,  
 RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
 RA FRITZ C., FUJITA M., FUJITA Y., FUNA S., GALIZZI A., GALLERON N.,  
 RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
 RA GUISEPI G., GUY B.J., HAGA K., HALECH J., HARWOOD C.R., HENAUT A.,  
 RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
 RA JORIS B., KARAMATA D., KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C.,  
 RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
 RA KURITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
 RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
 RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NOBACK M.,  
 RA NEONE D., O'REILLY M., OGAWA K., OGIIWARA A., OUDEGA B., PARK S.H.,  
 RA PARRO V., POHL T.M., PORTETELLE D., PORWOLLIK S., PRESCOTT A.M.,  
 RA PRESECAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
 RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADATE Y.,  
 RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOFFONE F.,  
 RA SEKIGUCHI J., SEKOWSKA A., SEROR S.J., SERROR P., SHIN B.S., SOLDI B.,  
 RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
 RA TAKEUCHI M., TAMAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
 RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
 RA VIARI A., WAMBUIT R., WEDLER E., WEDLER H., WEITZENEGER T.,  
 RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
 RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL NATURE 390:249-256(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;

RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF027868; G2619028; -  
DR EMBL: Z99114; E1185341; -  
SQ SEQUENCE 251 AA; 28855 MW; EE2C2BD7 CRC32;

Query Match 50.8%; Score 60; DB 2; Length 251;  
Best Local Similarity 38.5%; Pred. No. 3.18e-01;  
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;  
Db 5 EITKDNIEDNHIC 17  
|||:::|::|  
QY 1 EINKDDVQHKQVC 13

RESULT 6 PRELIMINARY; PRT; 309 AA.  
ID O06036  
AC O06036;  
DT 01-JUL-1997 (TREMREL. 04, CREATED)  
DT 01-JUL-1997 (TREMREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE EFSH.  
GN EPSH.  
OS LACTOCOCCUS LACTIS CREMORIS.  
OG PLASMID PN24000.  
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
OC LACTOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 97303093.  
RA VAN KRANENBURG R., MARUGG J.D., VAN SWAM I.I., WILLEM N.J., VOS W.M.;  
RT "Molecular characterization of the plasmid-encoded eps gene cluster  
essential for exopolysaccharide biosynthesis in Lactococcus lactis.";  
RL MOL. MICROBIOL. 24:387-397(1997).  
DR EMBL: U93364; G2072445; -  
KW PLASMID.  
SQ SEQUENCE 309 AA; 36818 MW; BC9E87D4 CRC32;

Query Match 50.08%; Score 59; DB 2; Length 309;  
Best Local Similarity 53.3%; Pred. No. 5.20e-01;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 206 EIKKNVNVKVICFS 220  
||:|:| |::|  
QY 1 EINKDDVQHKQVCFS 15

RESULT 7 PRELIMINARY; PRT; 150 AA.  
ID Q14290  
AC Q14290;  
DT 01-NOV-1996 (TREMREL. 01, CREATED)  
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)  
DE CAK BETA (CELL ADHESION KINASE BETA) (FRAGMENT).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE-BRAIN (HIPPOCAMPUS);  
RX MEDLINE: 95403356.  
RA SASAKI H., NAGURA K., ISHINO M., TOBIOKA H., KOTANI K., SASAKI T.;  
RT "Cloning and characterization of cell adhesion kinase beta, a novel  
protein-tyrosine kinase of the focal adhesion kinase subfamily.";  
RL J. BIOL. CHEM. 270:21206-21219(1995).  
DR EMBL: D45853; D1008884; -  
FT NON\_TER 150  
SQ SEQUENCE 150 AA; 17377 MW; 9B67D0A2 CRC32;

Query Match 49.2%; Score 58; DB 4; Length 150;  
Best Local Similarity 42.9%; Pred. No. 8.47e-01;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
Db 32 DVEKEDVRILKVCFS 45

QY 1 EINKDDVQHKQVC 14  
:::|::|::|::|

RESULT 8 PRELIMINARY; PRT; 317 AA.  
ID O44889  
AC O44889;  
DT 01-JUN-1998 (TREMREL. 06, CREATED)  
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE T20H9.3 PROTEIN.  
GN T20H9.3.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL NATURE 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RX BLANCHARD M., TWYMAN B.;  
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RA WATERSTON R.;  
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF040657; G2746901; -  
SQ SEQUENCE 317 AA; 37249 MW; 1687B237 CRC32;

Query Match 49.2%; Score 58; DB 5; Length 317;  
Best Local Similarity 33.3%; Pred. No. 8.47e-01;  
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 16 VNRRVDKTEICFDR 30  
:::|:|:|::|  
QY 2 EINKDDVQHKQVCFSR 16

RESULT 9 PRELIMINARY; PRT; 501 AA.  
ID O44865  
AC O44865;  
DT 01-JUN-1998 (TREMREL. 06, CREATED)  
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE K05F6.7 PROTEIN.  
GN K05F6.7.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-BRISTOL N2;  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL NATURE 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA DU Z., GOELA D.;  
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF040653; G2746872; -;  
 DR PROSITE; PS00298; HSP90; 1.  
 SQ SEQUENCE 501 AA; 58223 MW; 8C6B8F44 CRC32;

Query Match 49.2%; Score 58; DB 5; Length 501;  
 Best Local Similarity 60.0%; Pred. No. 8.47e-01;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 191 NVHSROVCF5 200  
 QY 6 DVQHKQVCF5 15  
 I::: |||||

RESULT 10  
 ID O44760 PRELIMINARY; PRT; 631 AA.  
 AC O44760;  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE T12F5.5 PROTEIN.  
 GN T12F5.5.  
 OS CAENORHABDITIS ELEGANS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECRETINTEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KRISTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL NATURE 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA CLARKE K., WOHLDMANN P., ROHLFING T., BAUER C.;  
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA WATERSTON R.;  
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; AF039718; G2773235; -;  
 SQ SEQUENCE 631 AA; 70685 MW; BE17D0F1 CRC32;

Query Match 47.5%; Score 56; DB 5; Length 631;  
 Best Local Similarity 50.0%; Pred. No. 2.21e+00;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 177 KNEITQKQVC 186  
 QY 4 KDDVQHKQVC 13  
 I::: |||||

RESULT 11  
 ID Q53922 PRELIMINARY; PRT; 692 AA.  
 AC Q53922; Q53948;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT GLCB.  
 GN GLCB.  
 OS STAPHYLOCOCCUS CARNOSUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 CC STAPHYLOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96180655.  
 RA CHRISTIANSEN I., HENGSTENBERG W.;  
 RT "Cloning and sequencing of two genes from Staphylococcus carnosus  
 RT coding for glucose-specific PTS and their expression in Escherichia  
 RT coli K-12.";  
 RL MOL. GEN. GENET. 250:375-379(1996).  
 CC -1- FUNCTION: THIS IS A COMPONENT OF THE  
 CC PHOSPHOENOLPYRUVATE-DEPENDENT SUGAR PHOSPHOTRANSFERASE SYSTEM  
 CC (PTS), A MAJOR CARBOHYDRATE ACTIVE -TRANSPORT SYSTEM.  
 CC THE IIIC DOMAINS CONTAIN THE SUGAR BINDING SITE AND THE  
 CC TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY  
 CC PHOSPHORYLATION SITE  
 CC (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS PHOSPHORYL GROUP TO  
 CC THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO THE SUGAR.  
 CC -1- CATALYTIC ACTIVITY: PROTEIN N-PHOSPHOTRANSFERASE.  
 CC HISTIDINE + SUGAR PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS A PTS EIIA DOMAIN.  
 CC -1- SIMILARITY: CONTAINS A PTS EIIB DOMAIN.  
 CC -1- SIMILARITY: CONTAINS A PTS EIIC DOMAIN.  
 DR EMBL; X93360; G1072419; -;  
 DR PFAM; PF00358; PTS-EIIA\_1; 1.  
 DR PFAM; PF00367; PTS-EIIB; 1.  
 KW PHOSPHOTRANSFERASE SYSTEM; SUGAR TRANSPORT; TRANSFERASE;  
 KW PHOSPHORYLATION; TRANSMEMBRANE.  
 SQ SEQUENCE 692 AA; 75343 MW; 6C24D6FB CRC32;

Query Match 47.5%; Score 56; DB 2; Length 692;  
 Best Local Similarity 42.9%; Pred. No. 2.21e+00;  
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 453 NKDNKHKLDACITR 466  
 QY 3 NKDDVQHKQVCF5R 16  
 I::: |::|

RESULT 12  
 ID O52118 PRELIMINARY; PRT; 136 AA.  
 AC O52118;  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE RFSA HOMOLOG (FRAGMENT).  
 OS BRUCELLA ABORIENS.  
 OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
 CC BRUCELLA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S2308;  
 RA ALLEN C.A., FICHT T.A.;  
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.



```

DR EMBL: AF021922; G2738957; -.
FT NON_TER 1 136
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 15369 MW; 9CACA BC0 CRC32:

Query Match 46.6%; Score 55; DB 2; Length 136;
Best Local Similarity 33.3%; Pred. No. 3.54e+00;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Db 45 IKRNRKRDHVCFFR 59
      |:::|:|:|
QY 2 INKDDVQHVKQVCFR 16

RESULT 13
ID O18068 PRELIMINARY; PRT; 347 AA.
AC O18068;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE T07G12.12 PROTEIN.
GN T07G12.12.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA CUMMINGS P.;
RA SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL NATURE 368:32-38(1994).
DR EMBL: Z82282; E1349307; -.
SQ SEQUENCE 347 AA; 39420 MW; 67C96E63 CRC32;

Query Match 46.6%; Score 55; DB 5; Length 347;
Best Local Similarity 46.7%; Pred. No. 3.54e+00;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 232 EINKDEIHNIQCHFP 246
      |:::|:|:|
QY 1 EINKDDVQHVKQVCFR 15

RESULT 14
ID O81143 PRELIMINARY; PRT; 467 AA.
AC O81143;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1999 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ACC SYNTHASE.
GN LE-ACS7.
OS LYCOPERICON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; SOLANACEAE; SOLANALES; SOLANACEAE; SOLANUM.
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-SOL3.
RX MEDLINE: 98374353.

```

---

```

RA SHIU O.Y., OETIKER J.H., YIP W.K., YANG S.F.;
RT "The promoter of LE-ACS7, an early flooding-induced
RT 1-aminocyclopropane-1-carboxylate synthase gene of the tomato, is
RT tagged by a Sol3 transposon.";
RL PROC. NATL. ACAD. SCI. U.S.A. 95:10334-10339(1998).
CC -|- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES
DR EMBL: AF043122; G3435102; -.
DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW PYRIDOXAL PHOSPHATE.
SQ SEQUENCE 467 AA; 53133 MW; F06E21C3 CRC32;

Query Match 46.6%; Score 55; DB 10; Length 467;
Best Local Similarity 50.0%; Pred. No. 3.54e+00;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 434 NKDDIQNQOCQSNK 447
      |:::|:|:|
QY 3 NKDDVQHVKQVCFR 16

RESULT 15
ID O16252 PRELIMINARY; PRT; 473 AA.
AC O16252;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE F44E7.7 PROTEIN.
GN F44E7.7.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA DU Z., SCHEET P., DELEHAUTY A.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF016421; G2291184; -.
DR PFAM; PF00083; sugar_tr; 1.
SQ SEQUENCE 473 AA; 52776 MW; 48F951EC CRC32;

Query Match 46.6%; Score 55; DB 5; Length 473;
Best Local Similarity 30.8%; Pred. No. 3.54e+00;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 265 KESVYPYRAICFDK 277
      |:|:|:|
QY 4 KDDVQHVKQVCFR 16

```

Search completed: Mon Aug 23 13:20:39 1999  
Job time : 19 secs.

---



```

RESULT 2
ID W46879 standard; Protein: 228 AA.
AC W46879;
DE 22-JUN-1998 (first entry)
DE Protein sequence encoded by a colon-specific gene.
KW Colon-specific gene; probe; detection; expression; human;
OS Homo sapiens.
PN US5733748-A.
PD 31-MAR-1998.
PF 06-JUN-1995; 469667.
PR 06-JUN-1995; US-469667.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen C, Yu G;
DR WPI: 98-229823/20.
DR N-PSDB; V16672.
PT Colon-specific nucleic acids - useful as probes for detecting colon
PT cancer micrometastases
PS Claim 1; Fig 5; 51pp; English.
CC W46876-80 and W4682-85 represent proteins encoded by colon-specific
CC genes. The polynucleotides encoding these proteins can be used
CC as probes to detect expression of the corresponding human genes,
CC e.g. in diagnostic assays for detecting micrometastases of colon cancer.
CC Recombinant cells containing the polynucleotides can be used to
CC produce the proteins, in order that antibodies can be raised and
CC used in further screening or diagnostics.
SQ Sequence 228 AA;

Query Match 100.0%; Score 263; DB 29; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.65e-20;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 wienprrpeinkddvqhkvcsfsgg 79
|||
QY 1 WIENDEIQWNPPEINKDDVQHKKVCFSTSGG 35
|||||

RESULT 3
ID R75704 standard; Protein: 951 AA.
AC R75704;
DE 11-NOV-1995 (first entry)
DE Eph-related CEK6.
KW CEK6; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW prognosis.
OS Gallus sp.
FH Key Location/Qualifiers
FT domain 426..444
FT /label= Extracellular_domain
FT WO9515375-A.
PN 08-JUN-1995.
PD 07-SEP-1994; U10140.
PR 03-DEC-1993; US-162809.
PA (LJOL-) LA JOLLA CANCER RES FOUND.
PI Pasquale EB, Sajjadi FG;
DR WPI: 95-215256/28.
DR N-PSDB; Q90652.
PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT cancer.
PS Claim 12; Page 37-41; 129pp; English.
CC Novel Eph-related PTK CEK6 cDNA clones (Q90652) were isolated from
CC chick embryo and embryonic brain cDNA libraries in phage lambda gt11.
CC The encoded CEK6 protein (R75704) is closely related to rat Elk.
CC CEK5 (R75712) and CEK10 (R75709). CEK6 transcripts were found in
CC 10-day embryos and in adult brain, lung, heart and skeletal muscle.
SQ Sequence 951 AA;

Query Match 31.2%; Score 82; DB 13; Length 951;
Best Local Similarity 47.6%; Pred. No. 2.17e-00;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 288 lewnprrtgrddvtnivc 308
:::||||| :||| :|||
```

```

QY 7 IQWNPPEINKDDVQHKKVC 27

RESULT 4
ID W27148 standard; Protein: 833 AA.
AC W27148;
DE 04-DEC-1997 (first entry)
DE HMG-CoA reductase degradation protein Hrd3p.
KW 3-hydroxy-3-methylglutaryl; coenzyme A; cholesterol; Hrd3p;
KW Hrd3p; hypercholesterolaemia; yeast.
OS Saccharomyces cerevisiae.
FH Key Location/Qualifiers
FT peptide 1..18
FT /note= "Putative signal sequence"
FT region 768..789
FT /note= "Strongly hydrophobic"
FT WO9707219-A2.
PN 27-FEB-1997.
PD 16-AUG-1996; IB1161.
PR 17-AUG-1995; US-002381.
PA (REGC ) UNIV CALIFORNIA.
PI Hampton R, Rine JD;
DR WPI: 97-165303/15.
DR N-PSDB; T85288.
PT 3-Hydroxy-3-methylglutaryl CoA reductase degradation polypeptide(s)
PT - useful as therapeutic agents to reduce hypercholesterolaemia
PS Claim 4; Page 85-87; 132pp; English.
CC The present sequence represents the 3-hydroxy-3-methylglutaryl (HMG)-CoA
CC reductase degradation (HRD) protein Hrd3p. Hrd proteins can be used to
CC regulate the degradation of HMG-CoA reductase, e.g. as therapeutic
CC agents to reduce hypercholesterolaemia, and to elucidate how the
CC cholesterol pathway modulates the degradation of HMG-CoA reductase. In
CC addition, as a result of their ability to bind the proteasome complex,
CC antibodies that specifically bind Hrd polypeptides can be used to
CC isolate the proteasome complex. Further, they can be used in various
CC assays to identify compounds that modify the degradation of HMG-CoA
CC reductase independently of the beneficial LDL receptor control axis.
CC The nucleic acid molecules can be used as molecular probes for the
CC isolation of homologous nucleic acid molecules and for the detection of
CC HRD nucleic acid molecules in yeast.
SQ Sequence 833 AA;

Query Match 27.8%; Score 73; DB 24; Length 833;
Best Local Similarity 33.3%; Pred. No. 1.53e-01;
Matches 8; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Db 684 witrkvnvwpssplnpndtqh 707
|| : : ||:| :|||
QY 1 WIENDEIQWNPPEINKDDVQH 23

RESULT 5
ID R44513 standard; Protein: 984 AA.
AC R44513;
DE 16-JUN-1994 (first entry)
DE elk.
KW Lambda gt11; expression vector; lambda-B1-Elk; protein tyrosine kinase;
KW Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;
KW phosphorylation; phosphorylated kinase insert domain; growth factor;
KW receptor kinase; platelet-derived growth factor receptor.
OS Rattus rattus.
FH Key Location/Qualifiers
FT peptide 1..17
FT /note= "Signal peptide"
FT misc_difference 61 /note= "Cysteine residue"
FT misc_difference 96 /note= "Cysteine residue"
FT misc_difference 106 /note= "Cysteine residue"
FT misc_difference 183 /note= "Cysteine residue"
FT misc_difference 196 /note= "Cysteine residue"
```

FT misc\_difference /note= "Cysteine residue"  
FT 225 /note= "Cysteine residue"  
FT misc\_difference 240 /note= "Cysteine residue"  
FT 240 /note= "Cysteine residue"  
FT misc\_difference 253 /note= "Cysteine residue"  
FT 253 /note= "Cysteine residue"  
FT misc\_difference 255 /note= "Cysteine residue"  
FT 255 /note= "Cysteine residue"  
FT misc\_difference 267 /note= "Cysteine residue"  
FT 267 /note= "Cysteine residue"  
FT misc\_difference 270 /note= "Cysteine residue"  
FT 270 /note= "Cysteine residue"  
FT misc\_difference 284 /note= "Cysteine residue"  
FT 284 /note= "Cysteine residue"  
FT misc\_difference 287 /note= "Cysteine residue"  
FT 287 /note= "Cysteine residue"  
FT misc\_difference 301 /note= "Cysteine residue"  
FT 301 /note= "Cysteine residue"  
FT misc\_difference 303 /note= "Cysteine residue"  
FT 303 /note= "Cysteine residue"  
FT misc\_difference 319 /note= "Cysteine residue"  
FT 319 /note= "Cysteine residue"  
FT misc\_difference 360 /note= "Cysteine residue"  
FT 360 /note= "Cysteine residue"  
FT misc\_difference 363 /note= "Cysteine residue"  
FT 363 /note= "Cysteine residue"  
FT misc\_difference 370 /note= "Cysteine residue"  
FT 370 /note= "Cysteine residue"  
FT misc\_difference 373 /note= "Cysteine residue"  
FT 373 /note= "Cysteine residue"  
FT modified\_site 425..427 /note= "N-glycosylation site"  
FT modified\_site 480..482 /note= "N-glycosylation site"  
FT CA2083521-A.  
PN 01-OCT-1993.  
PD 23-NOV-1992; 083521.  
PF 31-MAR-1992; US-861390.  
PR (MOUN ) MOUNT SINAI HOSPITAL CORP.  
PA Letwin K, Pawson A, Reedijk M;  
PI WPI: 93-406300/51.  
DR WPI: 0753471  
DR N-PSDB: 0753471  
PT Expression of phosphorylated exogenous protein - in host cells  
PT transformed with two vectors, one for the protein, the other for  
PT catalytic domain of protein kinase  
PS Disclosure; Fig 3; 55pp; English.  
PS This sequence is encoded by the elk cDNA and represents the protein  
CC tyrosine kinase, Elk. The Elk gene, B1, encode a protein which is  
CC a member of the Eph subfamily of protein tyrosine kinases. The Elk  
CC product is very similar to two other receptor-like tyrosine kinases,  
CC eph and eck. Lambda-B1-Elk may be used in the production of  
CC phosphorylated exogenous protein along with a further vector encoding  
CC the desired exogenous protein. These plasmid may be used to produce  
CC phosphorylated proteins in host cells which have no intrinsic capacity  
CC for phosphorylation, eg. bacteria. The system may be used for the  
CC expression of the phosphorylated kinase insert domain of a growth  
CC factor receptor kinase eg. platelet-derived growth factor receptor.  
SQ Sequence 984 AA;  
  
Query Match 27.8%; Score 73; DB 8; Length 984;  
Best Local Similarity 38.1%; Pred. No. 1.53e+01;  
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
  
Db 340 lewppppretggddvtniic 360  
: : l l l l : : l l : : l  
QY 7 IQNPPRPPEINKDDVQHKQVC 27  
  
RESULT 6  
ID R74635 standard; Protein; 459 AA.  
AC R74635;  
DT 09-NOV-1995 (first entry)  
  
Query Match 27.0%; Score 71; DB 13; Length 1487;  
Best Local Similarity 26.3%; Pred. No. 2.34e+01;  
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

DE Bacillus circulans CFTase partial amino acid sequence.  
KW Cycloinuloooligosaccharide fructanotransferase; Bacillus circulans;  
KW primer; PCR; amplification; fructose oligosaccharide; insulin.  
KW cyclic inulo-oligosaccharide fructanotransferase; insulin.  
OS Bacillus circulans.  
PN J07041500-A.  
PD 10-FEB-1995.  
PF 29-JUL-1993; 188147.  
PR 29-JUL-1993; JP-188147.  
PA (MITU ) MITSUBISHI KASEI CORP.  
DR WPI: 95-118702/16.  
DR N-PSDB; Q88184.  
PT Cyclo:inulo-oligosaccharide fructano:transferase and its DNA -  
PT useful for production of cyclic inulo-oligosaccharide from  
PT beta-2,1-fructose oligosaccharide, inulin, etc.  
PS Claim 3; Page 11-13; 16pp; Japanese.  
CC The partial amino acid sequence of the novel cycloinuloooligosaccharide  
CC fructanotransferase (CFTase) isolated from Bacillus circulans. The DNA  
CC sequence was isolated by PCR using the primers Q88186-7 which were  
CC generated based on the amino acid sequence of the purified protein  
CC (R74634). The resulting fragment was used as probe to obtain the  
CC complete DNA sequence (Q88185). The protein or bacteria producing the  
CC fructanotransferase from beta-2,1-fructose oligosaccharide, 1-fructose  
CC oligosaccharide and insulin.  
SQ Sequence 459 AA;  
  
Query Match 27.0%; Score 71; DB 13; Length 459;  
Best Local Similarity 26.3%; Pred. No. 2.34e+01;  
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
  
Db 19 wvsddmnhvnpvpalape 37  
: : l : : l : :  
QY 1 WIENDEIQWNPPEINKD 19  
  
RESULT 7  
ID R74634 standard; protein; 1487 AA.  
AC R74634;  
DT 09-NOV-1995 (first entry)  
DE Bacillus circulans CFTase protein.  
KW Cycloinuloooligosaccharide fructanotransferase; Bacillus circulans;  
KW primer; PCR; amplification; fructose oligosaccharide;  
KW cyclic inulo-oligosaccharide fructanotransferase; insulin.  
OS Bacillus circulans.  
FH Key Location/Qualifiers  
FT misc\_difference 705..720 /note= "sequence given in the specification is  
FT missing amino acids 705-720"  
J07041500-A.  
PD 10-FEB-1995.  
PF 29-JUL-1993; 188147.  
PR 29-JUL-1993; JP-188147.  
PA (MITU ) MITSUBISHI KASEI CORP.  
DR WPI: 95-118702/16.  
PT Cyclo:inulo-oligosaccharide fructano:transferase and its DNA -  
PT useful for production of cyclic inulo-oligosaccharide from  
PT beta-2,1-fructose oligosaccharide, inulin, etc.  
PS Claim 3; Page 7-11; 16pp; Japanese.  
CC The amino acid sequence of the novel cycloinuloooligosaccharide  
CC fructanotransferase isolated from Bacillus circulans. The amino acid  
CC sequence of the purified protein was determined. The sequence was  
CC used to generate primers (Q8816-7) used to isolate a partial clone of  
CC the gene (see Q88184). Subsequently the complete sequence of the  
CC gene was isolated and determined (Q88185). The protein or bacteria  
CC producing the protein can be used to produce cyclic  
CC inulo-oligosaccharide fructanotransferase from beta-2,1-fructose  
CC oligosaccharide, 1-fructose oligosaccharide and insulin.  
SQ Sequence 1487 AA;  
  
Query Match 27.0%; Score 71; DB 13; Length 1487;  
Best Local Similarity 26.3%; Pred. No. 2.34e+01;  
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```

PT or prodn.
PS Claim 18; Page 46-49; 133pp; English.
CC 4 Novel human EPH-like receptor protein tyrosine kinases, HEK5, HEK7,
CC HEK8 and HEK11 (R95089-92), respectively, were identified following
CC isolation of their encoding cDNAs (T02946-49) from a human foetal brain
CC cDNA library. HEK5, HEK7 and HEK8 show extensive homology to the
CC catalytic domain of chicken EPH-like receptors Cek5, Cek7 and Cek8.
CC HEK11 shows no homology to any known EPH-like receptor. Recombinant
CC HEK receptors (or their soluble extracellular domains) are produced by
CC coexpression of encoding sequences in procarcaryotic or eucaryotic host
CC cells, and are used to produce antibodies (utilised in diagnostic
CC assays), or to identify and purify ligands for HEK receptors, or
CC therapeutically to modulate the activation of cell-associated
CC receptors. Soluble HEK5 receptor may affect primarily brain and
CC pancreatic cells.
SQ Sequence 970 AA;

Query Match 25 58; Score 67; DB 15; Length 970;
Best Local Similarity 28.68; Pred. No. 5.41e+01;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0

Db 328 lewtpprdsaggredlyniic 348
   ::|||::|:
QY 7 IQNPPRPPEINKDDVQHKQVC 27

RESULT 10
ID R97018 standard; Protein; 994 AA.
AC R97018;
DT 19-MAR-1996 (first entry)
DE Receptor tyrosine kinase (neural kinase).
KE Receptor tyrosine kinase; neural kinase; Nuk; axon;
KW axonogenesis; nerve disorder.
OS Mus musculus.
FH Key Location/Qualifiers
FT peptide 1..26
FT /label= Sig_peptide
FT region 27..548
FT /label= Extracellular_region
FT /note= "the extracellular region (amino acids
FT 27-548) includes an N-terminal Ig-like
FT domain and an Ig-like Nuk repeat"
FT region 330..420
FT /label= FNIII
FT /note= "fibronectin type III repeat"
FT region 444..534
FT /label= FNIII
FT /note= "fibronectin type III repeat"
FT region 549..574
FT /label= Transmembrane_region
FT region 575..994
FT /label= Cytoplasmic_region
FT domain 623..707
FT /label= Tyrosine-kinase_catalytic_domain
FT W09530326-AL.
PN 09-NOV-1995.
PD 28-APR-1995; CA0254.
PR 29-APR-1994; US-235407.
PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
PI Henkemeyer M, Letwin K, Pawson A;
PI WPI; 95-393299/50.
DR N-PSDB; T07308.
DR DNA encoding neural receptor tyrosine kinase - useful in gene
PT therapy of nerve disorders, and for diagnosis and identification of
PT therapeutic agents
PS Claim 1; Page 70-74; 103pp; English.
CC A novel receptor tyrosine kinase (R87018), designated neural kinase
CC (Nuk) (R87018), is encoded by cDNA (T07308) obtd. from a mouse
CC embryo library. Nuk is expressed in migrating axons and is
CC involved in cell-cell interactions and axonogenesis in development
CC of the nervous system. Nuk or its fragments (pref. amino acids
CC 26-548 or 601-994) are used to identify (antagonists of the
CC (activated) receptor tyrosine kinase as a means of treating nerve

```

CC disorders and damage, or to raise antibodies used to monitor axon  
 CC migration and nerve cell interactions.  
 SQ Sequence 994 AA;

Query Match 25.5%; Score 67; DB 15; Length 994;

Best Local Similarity 28.6%; Pred. No. 5.41e+01;

Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 350 lewtprrdsggredlyvniic 370

::: ||| : : : : :

QY 7 IQWNPPRPEINKDDVQHKQVC 27

RESULT 11

ID W26366 standard; Protein; 994 AA.

AC W26366; 1997 (first entry)

DE Mouse Nuk tyrosine kinase.

KW Nuk tyrosine kinase; Eph receptor tyrosine kinase;

KW signal transduction; axonogenesis; neurodegenerative disease;

KW Alzheimer's disease; Parkinson's disease; Huntington's disease;

KW multiple sclerosis; amyotrophic lateral sclerosis;

KW Wernicke's disease; nerve damage; trauma; ischaemia; stroke.

OS Mus musculus.

FH Key Location/Qualifiers

FT peptide 1..26

FT /label= Sig\_peptide

FT protein 27..994

FT /label= Mat\_protein

FT domain 27..548

FT /label= Extracellular\_domain

FT region 330..420

FT /label= FNIII

FT /note= "fibronectin type III repeat"

FT region 444..534

FT /label= FNIII

FT /note= "fibronectin type III repeat"

FT domain 549..574

FT /label= Transmembrane\_domain

FT domain 575..994

FT /label= Cytoplasmic\_domain

FT region 623..888

FT /label= Tyrosine-kinase\_region

PN W09714966-A1.

PD 24-APR-1997.

PF 10-OCT-1996; CA0679.

PR 13-OCT-1995; US-005518.

PA (MOUN ) MOUNT SINAI HOSPITAL CORP.

PI Henkemeyer M, Pawson A;

DR WPI: 97-245245/22.

DR N-PSDB; T84528

PT Activation of ligand regulatory pathways by Eph subfamily receptor

PT tyrosine kinases - for stimulating or inhibiting axonogenesis,

PT useful for treatment of e.g. neurodegenerative diseases such as

PT Alzheimer's or Parkinson's diseases

PS Disclosure: Fig 3: 55pp; English.

CC Murine Nuk tyrosine kinase is an Eph subfamily receptor tyrosine

CC kinase that is essential for formation of the medial tract of the

CC anterior commissure of the brain, and which appears to play a

CC role in the formation of the habenular interpeduncle tract. Its

CC amino acid sequence was deduced from cDNA clones (see T84528)

CC isolated from an embryo cDNA library. The extracellular domain

CC of Nuk was shown to be sufficient for formation of the medial

CC tract. Eph subfamily receptor tyrosine kinases (e.g. the Nuk

CC extracellular domain) can be used in claimed methods to: activate

CC a ligand regulatory pathway in a cell; identify substances able to

CC bind a ligand for an Eph subfamily receptor tyrosine kinase; and

CC to affect neuronal development or regeneration, especially the

CC stimulation or inhibition of axonogenesis, in a mammal. Activation

CC of the ligand regulatory pathway results in downstream activation

CC of a series of regulatory pathways in cells that control gene

CC expression, cell division, cytoskeletal architecture, cell

CC metabolism, cell migration and cell-cell interactions. Substances

CC which activate the ligand regulatory pathway may be used for  
 CC stimulating or inhibiting neuronal development regeneration and  
 CC axonal migration associated with neurodegenerative disease e.g.  
 CC Alzheimer's, Parkinson's or Huntington's diseases, multiple  
 CC sclerosis, amyotrophic lateral sclerosis, deficiency diseases such  
 CC as Wernicke's disease, peripheral nerve damage, trauma and  
 CC ischaemia resulting from stroke.  
 SQ Sequence 994 AA;

Query Match 25.5%; Score 67; DB 24; Length 994;

Best Local Similarity 28.6%; Pred. No. 5.41e+01;

Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 350 lewtprrdsggredlyvniic 370

::: ||| : : : : :

QY 7 IQWNPPRPEINKDDVQHKQVC 27

RESULT 12

ID R75712 standard; Protein; 995 AA.

AC R75712;

DT 11-NOV-1995 (first entry)

DE Eph-related PTK Cek5.

KW Cek5; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

KW prognosis.

OS Gallus sp.

PN W09515375-A.

PD 08-JUN-1995.

PF 07-SEP-1994; U10140.

PR 03-DEC-1993; US-162809.

PA (LJOL-) LA JOLLA CANCER RES FOUND.

PI Pasquale EB, Sajjadi FG;

DR WPI: 95-215256/28.

DR N-PSDB; Q90660.

PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing

PT cancer.

PS Disclosure: Page 92-96; 129pp; English.

CC Probes derived from the Eph-related PTKs Cek4 (Q90659) and Cek5

CC (Q90660) were used to isolate novel cDNA clones (Q90652-58,

CC Q90661-62) from chicken embryo and embryonic brain libraries.

SQ Sequence 995 AA;

Query Match 25.5%; Score 67; DB 13; Length 995;

Best Local Similarity 28.6%; Pred. No. 5.41e+01;

Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 350 lewtprrdsggredlyvniic 370

::: ||| : : : : :

QY 7 IQWNPPRPEINKDDVQHKQVC 27

RESULT 13

ID R75709 standard; Protein; 1011 AA.

AC R75709;

DT 11-NOV-1995 (first entry)

DE Eph-related PTK Cek5+.

KW Cek5+; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;

KW prognosis.

OS Gallus sp.

PN W09515375-A.

PD 08-JUN-1995.

PF 07-SEP-1994; U10140.

PR 03-DEC-1993; US-162809.

PA (LJOL-) LA JOLLA CANCER RES FOUND.

PI Pasquale EB, Sajjadi FG;

DR WPI: 95-215256/28.

DR N-PSDB; Q90657.

PT Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing

PT cancer.

PS Claim 11; Page 71-75; 129pp; English.

CC A cDNA clone encoding a novel variant of Eph-related PTK Cek5,

CC Cek5+ (Q90657), was isolated from a chick embryo library in

CC lambda gt11. Cek5+ protein (R75709) contains a 16-amino acid

CC insertion in the juxtamembrane domain, and be a result of  
 CC alternative splicing. Cek5+ is exclusively expressed in the CNS.  
 SQ Sequence 1011 AA;

Query Match 25.5%; Score 67; DB 13; Length 1011;  
 Best Local Similarity 28.8%; Pred. No. 5.41e-01;  
 Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 350 lewtprrdsggredlvynlic 370  
 : : l l l : : : : :  
 QY 7 IOWNPPRPEINKDDVQHKQVC 27

## RESULT 14

ID W22481 standard; Protein: 700 AA.  
 AC W22481;  
 DT 07-OCT-1997 (first entry)  
 DE Plasmodium ebl-2.  
 KW DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;  
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;  
 KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;  
 KW Plasmodium.  
 OS Plasmodium falciparum.  
 PN W09040766-A2.  
 PD 19-DEC-1996.  
 PF 07-JUN-1996; U09508.  
 PR 07-JUN-1995; US-487826.  
 RA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;  
 PI Wellem's TE;  
 DR WPI: 97-052231/05.  
 DR N-PSDB; T72896.  
 DR New malaria vaccines - contains cysteine-rich DBL family protein  
 PT binding domains homologous domains of the Duffy and sialic acid  
 PT binding proteins  
 PS Disclosure: Page 46-48; 96pp; English.  
 CC This sequence represents ebl-2 of Plasmodium. Ebl-2 belongs to  
 CC the Duffy binding like (DBL) family of genes which have homology to the  
 CC Duffy antigen binding protein (DABP) and sialic acid binding protein  
 CC (SABP) conserved regions (see T72889 and T72888 respectively). The var  
 CC family of genes modulate cytoadherence and antigenic variation of  
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding  
 CC protein (DABP) are soluble proteins that appear in the culture  
 CC supernatant after infected erythrocytes release merozoites. DABP and SABP  
 CC mediate the binding of merozoites and schizonts to the erythrocyte  
 CC surface. These proteins are necessary for erythrocyte invasion by the  
 CC parasite. This sequence can be used in the compositions of the invention.  
 CC The compositions are for the treatment and prevention of malaria, and  
 CC comprise either a nucleotide sequence or encoded polypeptide of the  
 CC var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of  
 CC genes having homology with conserved regions of DABP and SABP. The  
 CC compositions are used for the treatment and prevention of malaria. They  
 CC are also used in the preparation of vaccines for inducing a protective  
 CC immune response in a mammal to Plasmodium merozoites (especially  
 CC Plasmodium falciparum or Plasmodium vivax).  
 SQ Sequence 700 AA;

Query Match 25.1%; Score 66; DB 23; Length 700;  
 Best Local Similarity 31.3%; Pred. No. 6.65e-01;  
 Matches 10; Conservative 10; Mismatches 9; Indels 3; Gaps 3;

Db 534 vwekkqgewmaikdhfgkgtdivqkgllvifs 565  
 : : : : : : : : : : : : : :  
 QY 1 WIENDEIQNWPPRPEINKD-D-VQHKQ-VCFS 29

## RESULT 15

ID R70235 standard; Protein: 700 AA.  
 AC R70235;  
 DT 22-SEP-1995 (first entry)  
 DE P. falciparum EBL-e2  
 KW Erythrocyte binding ligand; EBL-e2; binding domain; malaria; therapy;  
 KW vaccine.

OS Plasmodium falciparum.  
 PN W09507353-A.  
 PD 16-MAR-1995.  
 PF 07-SEP-1994; U10230.  
 PR 10-SEP-1993; US-119677.  
 RA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;  
 PI Wellem's TE;  
 DR WPI: 95-123427/16.  
 DR N-PSDB; Q83528.  
 DR New erythrocyte binding domain polypeptide(s) - isolated from  
 PT Plasmodium binding proteins, used in diagnosis, treatment and  
 PT prevention of malaria  
 PS Disclosure: Page 56-57; 81pp; English.  
 CC Erythrocyte binding ligand (EBL) family genes were cloned from  
 CC P. falciparum chromosome 7 subsegment libraries constructed during  
 CC genetic studies of the chloroquine resistance locus. The 4 genes,  
 CC EBL-e1 (Q83526), E31a (Q83527), EBL-e2 (Q83528) and Proj3 (Q83529),  
 CC encode the proteins given in R70233-36, respectively. The binding  
 CC domains of such proteins can be expressed e.g. in E. coli, yeast,  
 CC mammalian, insect, and in vaccinia virus and adenovirus-infected  
 CC cells, and provide protection against P. falciparum.  
 SQ Sequence 700 AA;

Query Match 25.1%; Score 66; DB 13; Length 700;  
 Best Local Similarity 31.3%; Pred. No. 6.65e-01;  
 Matches 10; Conservative 10; Mismatches 9; Indels 3; Gaps 3;

Db 534 vwekkqgewmaikdhfgkgtdivqkgllvifs 565  
 : : : : : : : : : : : : : :  
 QY 1 WIENDEIQNWPPRPEINKD-D-VQHKQ-VCFS 29

Search completed: Mon Aug 23 13:19:33 1999  
 Job time : 24 secs.



(TM)

Result. No.	Query			ID	Description	Pred. No.
	No.	Score	Match			
1	130	49.4	794	3	JG0168	6.62e-11
2	93	35.4	913	2	T02171	6.06e-04
3	93	35.4	820	2	T02152	6.06e-04
4	93	35.4	905	2	T02205	6.06e-04
5	83	31.6	3660	1	S02041	3.12e-02
6	82	31.2	952	2	I50612	4.58e-02
7	82	31.2	985	2	I31672	4.58e-02
8	77	29.3	984	2	A39753	2.98e-01
9	75	28.5	271	2	D70915	6.17e-01
10	75	28.5	273	2	F1615	6.17e-01
11	73	27.8	833	2	S48558	1.27e-00
12	72	27.4	1369	2	S70173	1.80e+00
13	71	27.0	108	2	S21420	2.57e+00
14	71	27.0	174	2	S61134	2.57e+00
15	70	26.6	605	2	A36361	3.63e+00
16	70	26.6	668	1	A44276	3.63e+00
17	70	26.6	1488	2	C70984	3.63e+00
18	69	26.2	249	2	C70436	5.13e+00
19	69	26.2	511	2	J55880	5.13e+00
20	69	26.2	533	2	A56110	5.13e+00
21	69	26.2	600	2	T01259	5.13e+00
22	69	26.2	624	2	S74952	5.13e+00
23	69	26.2	672	1	UYPVW1	5.13e+00

```

Query Match          35.4%; Score 93; DB 2; Length 794;
Best Local Similarity 51.4%; Pred. No. 6.06e-04;
Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPREV-KDILAKAKIEDFSRLTSGG 743
QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSTSSGG 35

RESULT 3
ENTRY T02152 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02152
REFERENCE T02152
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.O.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel.
#accession T02152
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-820 #label ELB
#cross-references EMBL:AF001263; NID:g2623766; PID:g2623767
#experimental_source lung
SUMMARY #length 820 #molecular-weight 91464 #checksum 3104

Query Match          35.4%; Score 93; DB 2; Length 820;
Best Local Similarity 51.4%; Pred. No. 6.06e-04;
Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPREV-KDILAKAKIEDFSRLTSGG 743
QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSTSSGG 35

RESULT 4
ENTRY T02205 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02205
REFERENCE T02205
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.O.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel.
#accession T02205
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-905 #label ELB
#cross-references EMBL:AF001261; NID:g2623762; PID:g2623763
#experimental_source lung
SUMMARY #length 905 #molecular-weight 101005 #checksum 5557

Query Match          35.4%; Score 93; DB 2; Length 905;
Best Local Similarity 51.4%; Pred. No. 6.06e-04;
Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPREV-KDILAKAKIEDFSRLTSGG 743
QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSTSSGG 35

RESULT 5
ENTRY S02041 #type complete
TITLE dystrophin, muscle - chicken

```

---

```

ALTERNATE_NAMES
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 07-Sep-1990 #sequence_revision 27-Jun-1994 #text_change
ACCESSIONS S02041; S02013; S71487
REFERENCE S02041
#authors Lemaire, C.; Heilig, R.; Mandel, J.L.
#journal Nucleic Acids Res. (1988) 16:11815-11816
#title Nucleotide sequence of chicken dystrophin cDNA.
#cross-references MUID:89098331
#accession S02041
#status translation not shown
#molecule_type mRNA
#residues 1-3660 #label LEM
#cross-references EMBL:X13369; NID:g63369; PID:g63370
#note 1869-His, 1885-Arg, and sequences lacking 1171-Met were
also found

REFERENCE S02013
#authors Lemaire, C.; Heilig, R.; Mandel, J.L.
#journal EMBO J. (1988) 7:4157-4162
#title The chicken dystrophin cDNA: striking conservation of the
C-terminal coding and 3' untranslated regions between man
and chicken.
#cross-references MUID:89210800
#accession S02013
#status nucleic acid sequence not shown
#molecule_type mRNA
#residues 1-3573; 'HA', 3576-3660 #label LEM2
REFERENCE S09071
#authors Heilig, R.; Lemaire, C.; Mandel, J.L.
#journal Nucleic Acids Res. (1987) 15:9129-9142
#title A 230kb cosmid walk in the Duchenne muscular dystrophy gene:
detection of a conserved sequence and of a possible
deletion prone region.
#cross-references MUID:88067745
#accession S71487
#molecule_type DNA
#residues 222-281 #label HEI
COMMENT Dystrophin is proposed to play a role in anchoring the cytoskeleton
to the plasma membrane.
COMMENT Defects in dystrophin are responsible for the Duchenne/Becker
muscular dystrophies.
CLASSIFICATION #superfamily dystrophin; alpha-actinin actin-binding domain
homology; spectrin/dystrophin repeat homology; WW repeat
homology
KEYWORDS actin binding; calmodulin binding; cytoskeleton; leucine
zipper; membrane-associated protein; muscle; structural
protein; tandem repeat; triple helix
FEATURE #domain alpha-actinin actin-binding domain homology
18-237 #label AAH\
253-327 #region hinge\
340-449 #domain spectrin/dystrophin repeat homology #label SN1\
450-558 #domain spectrin/dystrophin repeat homology #label SN2\
560-669 #domain spectrin/dystrophin repeat homology #label SN3\
670-719 #region hinge\
720-830 #domain spectrin/dystrophin repeat homology #label SN4\
838-936 #domain spectrin/dystrophin repeat homology #label SN5\
940-1047 #domain spectrin/dystrophin repeat homology #label SN6\
1049-1156 #domain spectrin/dystrophin repeat homology #label SN7\
1158-1265 #domain spectrin/dystrophin repeat homology #label SN8\
1267-1369 #domain spectrin/dystrophin repeat homology #label SN9\
1374-1479 #domain spectrin/dystrophin repeat homology #label S10\
1480-1570 #domain spectrin/dystrophin repeat homology #status
atypical #label S11\
1572-1678 #domain spectrin/dystrophin repeat homology #label S12\
1680-1784 #domain spectrin/dystrophin repeat homology #label S13\
1787-1877 #domain spectrin/dystrophin repeat homology #status
atypical #label S14\
1878-1984 #domain spectrin/dystrophin repeat homology #label S15\
1986-2103 #domain spectrin/dystrophin repeat homology #label S16\
2105-2211 #domain spectrin/dystrophin repeat homology #label S17\
2213-2319 #domain spectrin/dystrophin repeat homology #label S18\

```



```

Best Local Similarity 38.1%; Pred. No. 2.98e-01;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 340 LEWPPRETGGRDDVTYNIC 360
      :||:|||| :||| :|
QY 7 IQWPPRPPEINKDDVQHKVC 27

RESULT 9
ENTRY D70915 #type complete
TITLE Hypothetical protein Rv1433 - Mycobacterium tuberculosis
      (strain H37Rv)
ORGANISM #formal_name Mycobacterium tuberculosis
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
ACCESSIONS D70915
REFERENCE A70500
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
      C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry
      III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
      Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
      Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
      Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
      Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; K.;
      Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
      Skelton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
      Taylor, K.; Whitehead, S.; Barrell, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of Mycobacterium tuberculosis from
      the complete genome sequence.
#cross-references MUID:98295987
#accession D70915
##status preliminary; nucleic acid sequence not shown;
      translation not shown
##molecule_type DNA
##residues 1-271 #label COL
##cross-references GB:295844; GB:AL123456; NID:g3250713; PID:e318872;
      PID:g2131053
##experimental_source strain H37Rv

GENETICS
#gene Rv1433
SUMMARY #length 271 #molecular-weight 28701 #checksum 4925

Query Match 28.5%; Score 75; DB 2; Length 271;
Best Local Similarity 53.8%; Pred. No. 6.17e-01;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 90 WLDNDVVQVDPDR 102
      :||:|||| :|
QY 1 WIENDEIQWNP 13

RESULT 10
ENTRY C71615 #type complete
TITLE protein of the TMR7 family PFB0425c - malaria parasite
      (Plasmodium falciparum)
ORGANISM #formal_name Plasmodium falciparum
DATE 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change
ACCESSIONS C71615
REFERENCE A71600
#authors Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.;
      Aravind, L.; Koonin, E.V.; Shalton, S.; Mason, T.; Yu, K.;
      Fujii, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.;
      Lai, Z.; Schwartz, D.C.; Perte, M.; Salzberg, S.; Zhou,
      L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
      Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.L.
#journal Science (1998) 282:1126-1132
#title Chromosome 2 sequence of the human malaria parasite
      Plasmodium falciparum.
#accession C71615
##status preliminary; nucleic acid sequence not shown;
      translation not shown

##molecule_type DNA
##residues 1-273 #label GAR
##cross-references GB:AE001394; GB:AE001362; NID:g3845178; PID:g3845181;
      TIGR:PFB0425c
##experimental_source clone 3d7

GENETICS
#gene PFB0425c
SUMMARY #length 273 #molecular-weight 33327 #checksum 9328

Query Match 28.5%; Score 75; DB 2; Length 273;
Best Local Similarity 26.9%; Pred. No. 6.17e-01;
Matches 7; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

Db 136 WANQFDLSWNPKNKSEQHNVOEKDI 161
      | : ||| : : ||| :
QY 1 WIENDEIQWNPPEINKDDVQHKQV 26

RESULT 11
ENTRY S48558 #type complete
TITLE Probable membrane protein YLR207w - yeast (Saccharomyces
      cerevisiae)
ALTERNATE_NAMES hypothetical protein L8167.5
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change
ACCESSIONS S48558
REFERENCE S48545
#authors Pauley, A.
#submission submitted to the EMBL Data Library, September 1994
#description The sequence of S. cerevisiae cosmid 8167.
#accession S48558
##molecule_type DNA
##residues 1-833 #label PAU
##cross-references EMBL:U14913; NID:g544497; PID:g544502; MIPS:YLR207w

GENETICS
#gene SGD:HRD3
SUMMARY #cross-references SGD:S0004197; MIPS:YLR207w

KEYWORDS #map_position 12R
FEATURE transmembrane protein
2-18 #domain transmembrane #status predicted #label TM1\
768-784 #domain transmembrane #status predicted #label TM2
SUMMARY #length 833 #molecular-weight 95480 #checksum 417

Query Match 27.8%; Score 73; DB 2; Length 833;
Best Local Similarity 33.3%; Pred. No. 1.27e+00;
Matches 8; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Db 684 WITREKVNTWKSPSLNPNDTQH 707
      || : : ||| : |||
QY 1 WIENDEIQ-WNPPEINKDDVQVH 23

RESULT 12
ENTRY S70713 #type complete
TITLE protein-tyrosine kinase let-23 precursor homolog -
      Caenorhabditis vulgariis
ALTERNATE_NAMES receptor tyrosine kinase let-23 homolog
ORGANISM #formal_name Caenorhabditis vulgariis
DATE 21-Apr-1997 #sequence_revision 09-May-1997 #text_change
ACCESSIONS S70713
REFERENCE S70712
#authors Sakai, T.; Koga, M.; Ohshima, Y.
#journal J. Mol. Biol. (1996) 256:548-555
#title Genomic structure and 5' regulatory regions of the let-23
      gene in the nematode C. elegans.
#cross-references MUID:96177760
#accession S70713
##status nucleic acid sequence not shown
##molecule_type DNA
##residues 1-1369 #label SAK

```

```

##cross-references EMBL:D63427
GENETICS
#gene
#introns
421/1: 49/1; 83/1; 105/3; 155/3; 207/1; 280/1; 369/1; 408/1;
438/2; 555/1; 598/2; 673/2; 733/3; 830/3; 882/3; 1147/1;
1247/3; 1274/1; 1309/1
CLASSIFICATION
#superfamily unassigned Ser/Thr or Tyr-specific protein
kinases: protein kinase homology
KEYWORDS
ATP: phosphotransferase; transmembrane protein;
tyrosine-specific protein kinase
FEATURE
1-28
29-1369
#domain signal sequence #status predicted #label SIG\
#product protein-tyrosine kinase let-23 homolog #status
predicted #label MAT\
929-1194
#domain protein kinase homology #label KIN\
937-945
#region protein kinase ATP-binding motif
SUMMARY
#length 1369 #molecular-weight 156528 #checksum 7764
Query Match 27.4% Score 72; DB 2; Length 1369;
Best Local Similarity 36.7% Pred. No. 1.80e+00;
Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 2;
Db 121 WIEANEIQKWRNSTNQVDADIDYKTVNF 150
||| :||| : : :| :|
QY 1 WIEDEIQ-WNPPPEINKDDVQH-KQVCF 28
17-Mar-1999
RESULT 13
ENTRY S21420 #type complete
TITLE general stress protein homolog ytxJ - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
17-Mar-1999
ACCESSIONS S21420; F70003; S71003
REFERENCE
#authors Bolotin, A.P.; Khazak, V.E.; Ratmanova, K.I.; Yomantas, Y.I.;
Kozlov, Y.I.
#submission submitted to the EMBL Data Library, May 1992
#description Cloning and nucleotide structure of araA-arg gene of
Bacillus subtilis 168.
#accession S21420
##molecule_type DNA
##residues 1-108 #label BOL
##cross-references EMBL:X65945; NID:g39812; PID:g39815
##experimental_source strain 168
REFERENCE A69580
#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kashanara, Y.; Kiehr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Mauel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetille, D.; Porwolik, S.; Prescott,
A.M.; Pressecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon, E.;
Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,

```

```

B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vanderbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenecker, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references MUID:98044033
#accession F70003
#status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-108 #label KUN
##cross-references GB:299119; GB:AL009126; NID:g2635411; PID:el185849;
PID:g2635460
##experimental_source strain 168
REFERENCE S71000
#authors Varon, D.; Brody, M.S.; Price, C.W.
#journal Mol. Microbiol. (1996) 20:339-350
#title Bacillus subtilis operon under the dual control of the
general stress transcription factor sigma(B) and the
sporulation transcription factor sigma(H).
#cross-references MUID:96310371
#accession S71003
##molecule_type DNA
##residues 1-56 #label VAR
##cross-references EMBL:L31845; NID:g556013; PID:g556017
##experimental_source strain 168, substrain Marburg
GENETICS
#gene ytxJ
SUMMARY
#length 108 #molecular-weight 12402 #checksum 4574
Query Match 27.0% Score 71; DB 2; Length 108;
Best Local Similarity 22.7% Pred. No. 2.57e+00;
Matches 5; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
Db 84 IONGEVKWHSTHSQITEAAIEQ 105
||| :| :| :| :| :| :|
QY 2 IENDEIQWNPPEINKDDVQH 23
12-Dec-1997
RESULT 14
ENTRY S61134 #type complete
TITLE hypothetical protein YGL183c - yeast (Saccharomyces
cerevisiae)
ALTERNATE_NAMES hypothetical protein G1604
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
12-Dec-1997
ACCESSIONS S61134; S64200
REFERENCE S61128
#authors Bertani, I.; Cogliovina, M.; Zaccaria, P.; Klima, R.;
Bruschi, C.V.
#submission submitted to the EMBL Data Library, September 1995
#description The sequence analysis of a 7.9 kb DNA fragment from the left
arm of S.cerevisiae chromosome VII reveals five ORF's
including part of the CDC55, the entire COX4 and RPS26, and
two putative new genes.
#accession S61134
##molecule_type DNA
##residues 1-174 #label BER
##cross-references EMBL:X91489; NID:g1143557; PID:el99057; PID:g1143564
REFERENCE S64183
#authors Bruschi, C.V.; Cogliovina, M.; Bertani, I.; Klima, R.;
Zaccaria, P.; Delneri, D.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64200
##molecule_type DNA
##residues 1-174 #label BRU
##cross-references EMBL:Z72705; NID:g1322796; PID:e243495; PID:g1322797;

```

```

MIPS:YGL183C
##experimental_source strain S288C
GENETICS
  #map_position 7L
SUMMARY
  #length 174 #molecular-weight 20767 #checksum 5672

  Query Match      27.0%; Score 71; DB 2; Length 174;
  Best Local Similarity 25.0%; Pred. No. 2.57e+00;
  Matches      7; Conservative 14; Mismatches 7; Indels 0; Gaps 0;

Db    110 ESRDAAKIQENKQIRLKVVHLEKTT 137
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Qy    5 DEIQWNPPEINKDDVQHKQVCFRST 32

RESULT 15
ENTRY      A36361      #type fragment
TITLE      glucose transport protein SGLT1 - pig (fragment)
ORGANISM   #formal_name Sus scrofa domestica #common_name domestic pig
DATE       28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change
           08-Sep-1997

ACCESSIONS A36361
REFERENCE   A36361
#authors   Ohta, T.; Isselbacher, K.J.; Rhoads, D.B.
#journal   Mol. Cell. Biol. (1990) 10:6491-6499
#title     Regulation of glucose transporters in LUC-PK-1 cells: effects
           of D-glucose and monosaccharides.
#cross-references MUID:91061757
#accession A36361
#status     preliminary
#molecule_type mRNA
#residues   1-605 #label OHT
#cross-references GB:M34044; NID:gl64674; PID:gl64675
CLASSIFICATION
#superfamily proline carrier protein
KEYWORDS    transmembrane protein
SUMMARY     #length 605 #checksum 4232

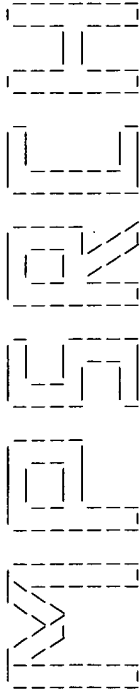
  Query Match      26.6%; Score 70; DB 2; Length 605;
  Best Local Similarity 34.5%; Pred. No. 3.63e+00;
  Matches      10; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

Db    519 EEDIO-EAPEETIEIEVPEEKKGCFRRT 546
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Qy    3 ENDEIWNPPREINKDDVQHKQVCFRST 31

Search completed: Mon Aug 23 13:18:49 1999
Job time : 13 secs.

```

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:17:33 1999; MasPar time 3.64 Seconds

Tabular output not generated. 272.041 Million cell updates/sec

Title: >US-09-049-696-48  
Description: (1-35) from US09049696.pep  
Perfect Score: 263  
Sequence: 1 WIENDEIOWNPRPEINKDDVOHKQVCFRTSSGG 35

Scoring table: PAM 150  
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss:prot37  
1:swissprot

Statistics: Mean 32.624; Variance 48.547; scale 0.672

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	98	37.3	903	1	ECLC_BOVIN EPITHELIAL CHLORIDE CH	7.90e-06
2	83	31.6	3660	1	DMD_CHICK DYSTROPHIN.	6.33e-03
3	77	29.3	984	1	EPB1_RAT EPHRIN TYPE-B RECEPTOR	7.75e-02
4	77	29.3	984	1	EPB1_HUMAN EPHRIN TYPE-B RECEPTOR	7.75e-02
5	71	27.0	108	1	YTXJ_BACSU HYPOTHETICAL 12.4 KD P	8.44e-01
6	71	27.0	174	1	YGT3_YEAST HYPOTHETICAL 20.8 KD P	8.44e-01
7	70	26.6	605	1	SL51_PIG SODIUM/GLUCOSE COTRANS	1.24e+00
8	70	26.6	668	1	VNCS_PAVH3 NONCAPSID PROTEIN NS-1	1.24e+00
9	69	26.2	672	1	VNCS_PAVH3 NONCAPSID PROTEIN NS-1	1.82e+00
10	69	26.2	672	1	VNCS_PAVH3 NONCAPSID PROTEIN NS-1	1.82e+00
11	68	25.9	413	1	YIHS_ECOLI HYPOTHETICAL 47.4 KD P	2.65e+00
12	68	25.9	672	1	VNCS_PAVH3 NONCAPSID PROTEIN NS-1	2.65e+00
13	68	25.9	691	1	VIVI_MAZE REGULATORY PROTEIN VIV	2.65e+00
14	68	25.9	2025	1	Y226_HUMAN TETRAPEPTIDE REPE	3.85e+00
15	67	25.5	94	1	SPC1_YEAST MICROSOMAL SIGNAL PEPT	3.85e+00
16	67	25.5	285	1	YDH5_SCHPO HYPOTHETICAL 32.8 KD P	3.85e+00
17	67	25.5	295	1	YDH2_XANAU HYPOTHETICAL PROTEIN I	3.85e+00
18	67	25.5	640	1	STYM_PODAN HYPOTHETICAL PROTEIN K	3.85e+00
19	67	25.5	751	1	Y226_HUMAN HYPOTHETICAL PROTEIN K	3.85e+00
20	67	25.5	986	1	EPB2_HUMAN EPHRIN TYPE-B RECEPTOR	3.85e+00
21	67	25.5	987	1	EPB2_COTJA EPHRIN TYPE-B RECEPTOR	3.85e+00
22	67	25.5	988	1	EPB2_CHICK EPHRIN TYPE-B RECEPTOR	3.85e+00
23	67	25.5	993	1	EPB2_MOUSE EPHRIN TYPE-B RECEPTOR	3.85e+00

24	66	25.1	479	1	YOS1_CAEEL HYPOTHETICAL 54.2 KD T	5.57e+00
25	66	25.1	746	1	CATA_MYCIT PEROXIDASE / CATALASE	5.57e+00
26	66	25.1	986	1	EPB4_CHICK EPHRIN TYPE-A RECEPTOR	5.57e+00
27	65	24.7	80	1	YIIF_ECOLI HYPOTHETICAL 9.2 KD PR	8.02e+00
28	65	24.7	205	1	CPCF_SYN2 PHYCOCYANOBILIN LYASE	8.02e+00
29	65	24.7	316	1	MIAA_ECOLI TRNA DELTA(2)-ISOPENTE	8.02e+00
30	65	24.7	652	1	HS70_LEIAM HEAT SHOCK 70 KD PROTE	8.02e+00
31	65	24.7	931	1	DAP1_YEAST DIPEPTIDYL AMINOPEPTID	8.02e+00
32	65	24.7	1131	1	CPSE_CHICK C-PROTEIN, SKELETAL MU	8.02e+00
33	64	24.3	71	1	NXL1_NAJNA LONG NEUROTOXIN 1 (TOX	1.15e+01
34	64	24.3	71	1	NXL2_NAJNA LONG NEUROTOXIN 2 (TOX	1.15e+01
35	64	24.3	71	1	NXL5_NAJNA LONG NEUROTOXIN 5 (TOX	1.15e+01
36	64	24.3	71	1	NXL4_NAJNA LONG NEUROTOXIN 4 (TOX	1.15e+01
37	64	24.3	210	1	CD8B_PONPY T-CELL SURFACE GLYCOPR	1.15e+01
38	64	24.3	243	1	YNO6_CAEEL HYPOTHETICAL 27.9 KD P	1.15e+01
39	64	24.3	288	1	YLM9_CAEEL HYPOTHETICAL 32.9 KD P	1.15e+01
40	64	24.3	546	1	CALX_SOYBN CALNEXIN HOMOLOG REPTO	1.15e+01
41	64	24.3	625	1	TPOR_MOUSE THROMBOPOIETIN PRECU	1.15e+01
42	64	24.3	1827	1	SUIS_HUMAN SUCRASE-ISOMALTASE, IN	1.15e+01
43	63	24.0	183	1	XJGA_ECOLI HYPOTHETICAL 21.4 KD P	1.64e+01
44	63	24.0	653	1	HS70_LEIDO HEAT SHOCK 70 KD PROTE	1.64e+01
45	63	24.0	998	1	EPB7_HUMAN EPHRIN TYPE-A RECEPTOR	1.64e+01

ALIGNMENTS

RESULT 1	ECLC_BOVIN	STANDARD;	PRT;	903 AA.
ID	P54281;			
AC	01-OCT-1996 (REL. 34, CREATED)			
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE CHANNEL).			
DE	BOS TAURUS (BOVINE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
OC	ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TRACHEA;			
RX	MEDLINE; 96125078.			
RA	CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,			
RA	ARRATE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;			
RT	"Cloning of an epithelial chloride channel from bovine trachea.;"			
RL	J. BIOL. CHEM. 270:31016-31026(1995).			
CC	-!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME; MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND TRANSEPITHELIAL TRANSPORT			
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
CC	-!- TISSUE SPECIFICITY: TRACHEA.			
CC	-!- PTM: PHOSPHORYLATED BY CAM-KINASE II.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; U36445; G1184066; ..			
DR	IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;			
KW	CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.			
FT	TRANSMEM 7 27 POTENTIAL.			
FT	TRANSMEM 331 351 POTENTIAL.			
FT	TRANSMEM 617 637 POTENTIAL.			
FT	TRANSMEM 883 903 POTENTIAL.			
FT	CARBOHYD 75 75 POTENTIAL.			
FT	CARBOHYD 278 278 POTENTIAL.			
FT	CARBOHYD 360 360 POTENTIAL.			
FT	CARBOHYD 372 372 POTENTIAL.			
FT	CARBOHYD 504 504 POTENTIAL.			

FT	DOMAIN	3052	3085	WW DOMAIN.
----	--------	------	------	------------

DR PFAM; PF00069; pkinase; 1.

```

FT DOMAIN 3086 3357 CYS-RICH.
FT VARIANT 1171 1171 MISSING.
FT VARIANT 1869 1869 Q -> H.
FT VARIANT 1885 1885 K -> R.
SQ SEQUENCE 3660 AA; 422874 MW; AF61A205 CRC32;

Query Match 31.6%; Score 83; DB 1; Length 3660;
Best Local Similarity 44.0%; Pred. No. 6.33e-03;
Matches 11; Conservative 6; Mismatches 7; Indels 1; Gaps 1;

Db 1610 DAEIANGKATRKETEKKROVOLKNIC 1634
QY 4 NDEIQW-NPPRPEINKDDVOHKQVC 27

RESULT 3
ID EPB1_RAT STANDARD; PRT: 984 AA.
AC P09759;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE EPHRIN TYPE-B RECEPTOR 1 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE KINASE RECEPTOR EPH-2) (ELK).
GN EPHB1 OR EPH2 OR ELK.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN 11]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RX MEDLINE; 91203869.
RA LHOTAK V., GREER P., LETWIN K., PAWSON T.;
RT "Characterization of elk, a brain-specific receptor tyrosine kinase.";
RN 12]
RP SEQUENCE OF 605-984 FROM N.A.
RC STRAIN=WISTAR; TISSUE=BRAIN;
RX MEDLINE; 91203869.
RA LETWIN K., YEE S.P., PAWSON T.;
RT "Characterization of elk, a brain-specific receptor tyrosine kinase.";
RN 13]
CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. BINDS TO
CC EPHRIN-B1, -B2 AND -B3. MAY BE INVOLVED IN CELL-CELL INTERACTIONS
CC IN THE NERVOUS SYSTEM.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: RESTRICTED TO BRAIN AND TESTES.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M59814; ; NOT_ANNOTATED_CDS.
CC EMBL; X13411; G56095; .
CC PIR; S04327; S04327.
CC PIR; A39753; A39753.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
CC PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
CC PROSITE; PS01186; EGF_2; UNKNOWN_1.
CC PFAM; PF00041; fn3; 2.
CC PFAM; PF00069; pkinase; 1.
CC PFAM; PF00536; SAM; 1.

```



```
DR HSP; P00523; 2PTK.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 984 EPHRIN TYPE-B RECEPTOR 1.
FT DOMAIN 18 540 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 541 563 POTENTIAL.
FT DOMAIN 564 984 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 18 318 CYS-RICH.
FT DOMAIN 319 429 FIBRONECTIN TYPE-III.
FT DOMAIN 430 527 FIBRONECTIN TYPE-III.
FT DOMAIN 619 882 PROTEIN KINASE.
FT NP_BIND 625 633 ATP (BY SIMILARITY).
FT BINDING 651 651 ATP (BY SIMILARITY).
FT ACT_SITE 744 744 BY SIMILARITY.
FT MOD_RES 778 778 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 334 334 POTENTIAL.
FT CARBOHYD 426 426 POTENTIAL.
FT CARBOHYD 480 480 POTENTIAL.
SQ SEQUENCE 984 AA; 109882 MW; 0E9F241B CRC32;

Query Match 29.3%; Score 77; DB 1; Length 984;
Best Local Similarity 38.1%; Pred. No. 7.75e-02;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 340 LEWHPRETGGRDVTYNIIC 360
QY 7 IQWNPPEINKDDVQHKQVC 27
:::|||||:::

RESULT 4
ID EPBL_HUMAN STANDARD; PRT; 984 AA.
AC P54762;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE EPHRIN TYPE-B RECEPTOR 1 PRECURSOR (EC 2.7.1.112) (TYROSINE-PROTEIN
DE KINASE RECEPTOR EPH-2) (NET).
GN EPB1 OR EPH2 OR NET.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
[1]
RN SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RX MEDLINE; 96115594.
RA TANG X.X., BIEGEL J.A., NYCUM L.M., YOSHIOKA A., BRODEUR G.M.,
RA PLEASURE D.E., IKEGAKI N.;
RT "cDNA cloning, molecular characterization, and chromosomal
RT localization of NEP(EPH2), a human EPH-related receptor
RT protein-tyrosine kinase gene preferentially expressed in brain.";
RL GENOMICS 29:426-437(1995).
CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-B FAMILY. BINDS TO
CC EPHRIN-B1, -B2 AND -B3. MAY BE INVOLVED IN CELL-CELL INTERACTIONS
CC IN THE NERVOUS SYSTEM.
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN. BELONGS TO THE EPHRIN RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L40636; G1100112; -.
DR MIM; 600600; -.
```

```
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00790; RECEPTOR_TYR_KIN_V_1; 1.
DR PROSITE; PS00791; RECEPTOR_TYR_KIN_V_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PFAM; PF00041; fn3; 2.
DR PFAM; PF00069; PKinase; 1.
DR PFAM; PF00536; SAM; 1.
DR HSP; P00523; 2PTK.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 984 EPHRIN TYPE-B RECEPTOR 1.
FT DOMAIN 18 540 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 541 563 POTENTIAL.
FT DOMAIN 564 984 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 18 318 CYS-RICH.
FT DOMAIN 319 429 FIBRONECTIN TYPE-III.
FT DOMAIN 430 527 FIBRONECTIN TYPE-III.
FT DOMAIN 619 882 PROTEIN KINASE.
FT NP_BIND 625 633 ATP (BY SIMILARITY).
FT BINDING 651 651 ATP (BY SIMILARITY).
FT ACT_SITE 744 744 BY SIMILARITY.
FT MOD_RES 778 778 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 334 334 POTENTIAL.
FT CARBOHYD 426 426 POTENTIAL.
FT CARBOHYD 480 480 POTENTIAL.
SQ SEQUENCE 984 AA; 109884 MW; BSAC581A CRC32;

Query Match 29.3%; Score 77; DB 1; Length 984;
Best Local Similarity 38.1%; Pred. No. 7.75e-02;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 340 LEWHPRETGGRDVTYNIIC 360
QY 7 IQWNPPEINKDDVQHKQVC 27
:::|||||:::

RESULT 5
ID YTXJ_BACSU STANDARD; PRT; 108 AA.
AC P39914;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 12.4 KD PROTEIN IN MURC-AROA INTERGENIC REGION (ORF2)
DE (ORF3).
GN YTXJ.
OS BACILLUS SUBTILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=168;
RA BOLOTIN A.P., KHAZAK V.E., RATMANOVA K.I., YOMANTAS Y.I.,
RA KOZLOV Y.I.;
RL SUBMITTED (MAY-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN SEQUENCE FROM N.A.
RC MEDLINE; 98048467.
RA LAPIDUS A., GALLERON N., SOROKIN A., EHRLICH S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb irnb-dnaB region.";
RL MICROBIOLOGY 143:3431-3441(1997).
[3]
RN SEQUENCE OF 1-56 FROM N.A.
RC STRAIN=168 / MARBURG;
RX MEDLINE; 96310371.
RA VARON D., BRODY M.S., PRICE C.W.;
RT "Bacillus subtilis operon under the dual control of the general
RT stress transcription factor sigma B and the sporulation transcription
RT factor sigma H.";
RL MOL. MICROBIOL. 20:339-350(1996).
```

AC P26429;

```
Query Match          26.6%; Score 70; DB 1; Length 605;  
Best Local Similarity 34.5%; Pred. No. 1.34e+00;  
Matches 10; Conservative 8; Mismatches 10; Indels 1; Gaps 1;
```

```

Query Match      26.6%; Score 70; DB 1: Length 605;
Best Local Similarity 34.5%; Pred No. 1,24e+00;
Matches 10; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

Db 519 EEEDIO-EAPEETIETIEVPEEKKGCFRT 546
      |::|||::|:|::|::|
QY 3 ENDEIQNPPRPEINKDDVOHKQVCFST 31

```

```
db      67 WQOSEDNEWESTVDENTKKQV   87
       |:::|::|:|::|:|
ov      1 WIENDEIOWNPPRPEINKDDV   21
```

CC  
CC -1- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION.  
CC  
CC -1- SIMILARITY: BELONGS TO THE PARVOVIRUSES NONCAPSID PROTEIN FAMILY.  
CC  
CC -----  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and/or commercial  
CC distribution without prior permission from the copyright holder is prohibited.

```

RESULT 11
ID YIHS_ECOLI STANDARD; PRT: 413 AA.
AC P32140.
DT 01-OCT-1993 (REL. 27, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 47.4 KD PROTEIN IN GLNA-RBN INTERGENIC REGION.
GN YIHS.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE: 93347969.
RA PLUNKETT G. III, BURLAND V.D., DANIELS D.L., BLATTNER F.R.:
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes."
RL NUCLEIC ACIDS RES. 21:3391-3398(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; LI9201; G304984; ALT_INIT.
DR EMBL; AF000464; G1790313; ALT_INIT.
DR PIR; S40824; S40824.
DR ECOGENE; EGI1845; YIHS.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 413 AA; 47432 MW; 70410A5F CRC32;

Query Match 25.9%; Score 68; DB 1; Length 413;
Best Local Similarity 47.4%; Pred. No. 2.65e+00;
Matches 9; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Db 221 DTQWNL-POYKNDPAHR 238
QY 6 ETQWNPPEINKDDVQHK 24

RESULT 12
ID VNCS_MUNIV STANDARD; PRT: 672 AA.
AC P03134;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE NONCAPSID PROTEIN NS-1 (NONSTRUCTURAL PROTEIN NS1) (NCVP1).
GN NS1.
OS MURINE MINUTE VIRUS (MURINE PARVOVIRUS).
OC VIRUSES; SSDNA VIRUSES; PARVOVIRIDAE; PARVOVIRINAE; PARVOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83143341.
RA ASTELL C.R., THOMSON M., MERCHINSKY M., WARD D.C.:
RT "The complete DNA sequence of minute virus of mice, an autonomous
RT parvovirus."
RL NUCLEIC ACIDS RES. 11:999-1018(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL; J02275; G332295; -
DR EMBL; V01115; G60912; ALT_INIT.
DR PIR; A03696; UYPVIM.
DR PFAM; PF01057; Parvo_Ns1; 1.
DR TRANSFAC; T02375; -
KW NONSTRUCTURAL PROTEIN; NONCAPSID PROTEIN; DNA REPLICATION;
KW ATP-BINDING.
FT NP_BIND 399 406 ATP (POTENTIAL).
SQ SEQUENCE 672 AA; 76248 MW; 74284AE0 CRC32;

Query Match 25.9%; Score 68; DB 1; Length 672;
Best Local Similarity 23.8%; Pred. No. 2.65e+00;
Matches 5; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Db 67 WQSEDMEWETTVDENKQV 87
QY 1 WIENDEIQWNPPEINKDDV 21

RESULT 13
ID VIVI_MAIZE STANDARD; PRT: 691 AA.
AC P26307;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN VIVIPAROUS-1.
GN VP1.
OS ZEA MAYS (MAIZE).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES;
OC POACEAE; ZEA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. W22;
RX MEDLINE: 91364172.
RA MCCARTY D.R., HATTORI T., CARSON C.B., VASIL V., LAZAR M., VASIL I.K.:
RT "The Viviparous-1 developmental gene of maize encodes a novel
RT transcriptional activator."
RL CELL 66:895-905(1991).
CC -----
CC FUNCTION: TRANSCRIPTIONAL ACTIVATOR SPECIFICALLY REQUIRED FOR
CC EXPRESSION OF THE MATURATION PROGRAM IN THE SEED DEVELOPMENT.
CC PROBABLY POTENTIATES THE RESPONSE TO THE SEED-SPECIFIC HORMONE
CC ABSCISIC ACID (ABA). MAY BIND TO DNA INDIRECTLY.
CC SUBCELLULAR LOCATION: CYTOPLASMIC OR NUCLEAR.
CC TISSUE SPECIFICITY: SEED.
CC DEVELOPMENTAL STAGE: BETWEEN 10 AND 30 DAYS AFTER POLLINATION.
CC SIMILARITY: TO A. THALIANA ABI3 AND RICE VPL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60214; G168605; -
DR PIR; A40024; A40024.
DR MAZEDB; 65585; -
DR TRANSFAC; T01087; -
KW DEVELOPMENTAL PROTEIN; TRANSCRIPTION REGULATION; DNA-BINDING;
KW ACTIVATOR.
FT DOMAIN 1 170 PRO/ASP/GLU/SER/THR-RICH (PEST DOMAIN)
FT DOMAIN 1 121 TRANSCRIPTIONAL ACTIVATION.
SQ SEQUENCE 691 AA; 73328 MW; ICEAFB92 CRC32;

Query Match 25.9%; Score 68; DB 1; Length 691;
Best Local Similarity 26.9%; Pred. No. 2.65e+00;
Matches 7; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Db 222 WYQNHLOKRPDYWEEAGLHVOL 247

```

QY 1 WIENDEIQWNPPEINKDDVQHKQV 26

RESULT 14  
ID TTC3\_HUMAN STANDARD; PRT: 2025 AA.  
AC P53804;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE TETRATRICOPTIDE REPEAT DOMAIN 3 (TPR REPEAT PROTEIN D).  
GN TTC3 OR TPRD.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE: 96281123.  
RA OHIRA M., OOTSUYAMA A., SUZUKI E., ICHIKAWA H., SEKI N.,  
RA NAGASE T., NOMURA N., OHKI M.;  
RT "Identification of a novel human gene containing the  
RT tetratricopeptide repeat domain from the Down syndrome region of  
RT chromosome 21.";  
RL DNA RES. 3:9-16(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=FETAL BRAIN;  
RX MEDLINE: 97103476.  
RA TSUKAHARA F., HATTORI M., MURAKI T., SAKAKI Y.;  
RT "Identification and cloning of a novel cDNA belonging to  
RT tetratricopeptide repeat gene family from Down syndrome-critical  
RT region 21q22.2.";  
RL J. BIOCHEM. 120:820-827(1996).  
RN [3]  
RP TISSUE SPECIFICITY: FOUND IN ALL TISSUES EXAMINED.  
CC -1- TISSUE SPECIFICITY: CONTAINS 3 TPR DOMAINS.  
CC -1- SIMILARITY: CONTAINS 3 TPR DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D83077; D1012436; -  
CC EMBL; D84294; G1632762; -  
CC FIM; 602259; -  
CC PFAM; PF00097; zf-C3HC4; 1.  
KW REPEAT; TPR DOMAIN.  
FT REPEAT 231 264 TPR 1.  
FT REPEAT 265 298 TPR 2.  
FT REPEAT 299 332 TPR 3.  
FT DOMAIN 453 456 POLY-SER.  
FT DOMAIN 1020 1029 POLY-LYS.  
FT DOMAIN 1899 1902 POLY-LYS.  
FT DOMAIN 1018 1029 ARG/LYS-RICH (BASIC).  
FT DOMAIN 1172 1185 ARG/LYS-RICH (BASIC).  
FT DOMAIN 1563 1579 ARG/LYS-RICH (BASIC).  
SQ SEQUENCE 2025 AA; 229889 MW; 1115CC96 CRC32;

Query Match 25.98; Score 58; DB 1; Length 2025;  
Best Local Similarity 37.0%; Pred. No. 2.65e+00;  
Matches 10; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Db 787 RMEEDLRSPKPKNEOKETVDNVQRC 813  
QY 1 WIENDEIQWNPPEINKDDVQHKQVC 27

RESULT 15  
ID SPC1\_YEAST STANDARD; PRT: 94 AA.  
AC P46965;

DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE MICROSOMAL SIGNAL PEPTIDASE 10.8 KD SUBUNIT (EC 3.4.-.-).  
GN SPC1 OR YJR0108W.  
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB320 / ATCC 37323;  
RX MEDLINE: 96279206.  
RA FANG H., PANZNER S., MULLINS C., HARTMANN E., GREEN M.;  
RT "The homologue of mammalian SPC12 is important for efficient signal  
RT peptidase activity in *Saccharomyces cerevisiae*.";  
RL J. BIOL. CHEM. 271:16460-16465(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
DE HAAN M., SMITS P.H.M., GRIVELL L.A.;  
RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- FUNCTION: PART OF THE SIGNAL PEPTIDASE COMPLEX (SPC), EXACT  
CC FUNCTION IS NOT KNOWN.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U26257; G1002498; -  
CC EMBL; Z49510; E318976; -  
CC EMBL; Z49511; E319018; -  
CC SCD; L0003006; SPC1.  
KW HYDROLASE; PROTEASE; MICROsome; ENDOPLASMIC RETICULUM; TRANSMEMBRANE.  
FT TRANSMEM 29 49 POTENTIAL.  
FT TRANSMEM 51 71 POTENTIAL.  
SQ SEQUENCE 94 AA; 10819 MW; 559490FF CRC32;  
Query Match 25.58; Score 67; DB 1; Length 94;  
Best Local Similarity 35.0%; Pred. No. 3.85e+00;  
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Db 73 WYKQKLRWAQPKIEINVDQ 92  
QY 1 WIENDEIQWNPPEINKDD 20

Search completed: Mon Aug 23 13:17:44 1999  
Job time : 11 secs.

THIS PAGE BLANK (uspro)



RT chloride channel.";  
 RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF052746; G3560547; -;  
 SQ SEQUENCE 901 AA; 100039 MW; E52BF02E CRC32;

Query Match 36.5%; Score 96; DB 11; Length 901;  
 Best Local Similarity 38.2%; Pred. No. 1.38e-04;  
 Matches 13; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Db 706 VENGKIVLNPPRDPVQEEAEIATVEDFNRTSGG 739  
 :||| | ||||| : : : : ||| :|||  
 QY 2 IENDEIQWNPPEINKDDVQHKQVCFSTSGG 35

RESULT 3  
 ID O18742 PRELIMINARY; PRT; 794 AA.

AC O18742;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLIA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;  
 RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
 RA GOODWIN A., PAULI B.U.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF001262; G2623765; -;  
 SQ SEQUENCE 794 AA; 88509 MW; 2D450EBB CRC32;

Query Match 35.4%; Score 93; DB 6; Length 794;  
 Best Local Similarity 51.4%; Pred. No. 4.91e-04;  
 Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPPPEV-KDDLAKAKIEDFSRLTSGG 743  
 :||| | ||||| : ||| :|||  
 QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSRTSGG 35

RESULT 4  
 ID O18743 PRELIMINARY; PRT; 820 AA.

AC O18743;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLIA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;  
 RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
 RA GOODWIN A., PAULI B.U.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF001263; G2623767; -;  
 SQ SEQUENCE 820 AA; 91464 MW; 171ED21C CRC32;

Query Match 35.4%; Score 93; DB 6; Length 820;  
 Best Local Similarity 51.4%; Pred. No. 4.91e-04;  
 Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPPPEV-KDDLAKAKIEDFSRLTSGG 743  
 :||| | ||||| : ||| :|||  
 QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSRTSGG 35

RESULT 5  
 ID O18741 PRELIMINARY; PRT; 905 AA.  
 AC O18741;

DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC ARTIODACTYLIA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;  
 RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
 RA GOODWIN A., PAULI B.U.;  
 RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: AF001261; G2623763; -;  
 DR PFAM: PF00041; fn3; 1.  
 SQ SEQUENCE 905 AA; 101005 MW; 86E65459 CRC32;

Query Match 35.4%; Score 93; DB 6; Length 905;  
 Best Local Similarity 51.4%; Pred. No. 4.91e-04;  
 Matches 18; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 710 VENGKIILNPPPEV-KDDLAKAKIEDFSRLTSGG 743  
 :||| | ||||| : ||| :|||  
 QY 2 IENDEIQWNPPEINKDDVQHKQV-CFSRTSGG 35

RESULT 6  
 ID O75189 PRELIMINARY; PRT; 1115 AA.

AC O75189;  
 DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE KIAA0710 PROTEIN.

GN KIAA0710.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
 OC CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN;  
 RX MEDLINE: 98403880.  
 RA ISHIKAWA K., NAGASE T., SUYAMA M., MIYAJIMA N., TANAKA A., KOTANI H.,  
 RA NOMURA N., OHARA O.;  
 RT "Prediction of the coding sequences of unidentified human genes. X.  
 The complete sequences of 100 new cDNA clones from brain which can  
 code for large proteins in vitro.";  
 RL DNA RES. 5:189-176(1998).  
 DR EMBL: AB014610; D1032646; -;  
 SQ SEQUENCE 1115 AA; 125105 MW; 2D27891E CRC32;

Query Match 33.1%; Score 87; DB 4; Length 1115;  
 Best Local Similarity 53.3%; Pred. No. 5.88e-03;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 827 WTDGDEMOWGPARAE 841  
 | : ||| :|||  
 QY 1 WIENDEIQWNPPE 15

RESULT 7  
 ID Q07494 PRELIMINARY; PRT; 952 AA.

AC Q07494;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE TYROSINE KINASE CEK6 RECEPTOR (EC 2.7.1.112) (FRAGMENT).  
 GN CEK6.

OS GALLUS GALLUS (CHICKEN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;



RX MEDLINE: 93288394.  
RA SAJJADI F.G., PASQUALE E.B.;  
RT "Five novel avian Eph-related tyrosine kinases are differentially  
expressed.";  
RL ONCOGENE 8:1807-1813(1993).  
CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.  
CC -1- PROBABLY AN IMPORTANT COMPONENT OF SIGNAL TRANSDUCTION PATHWAYS.  
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP + PROTEIN  
CC TYROSINE PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
CC DOMAIN. BELONGS TO THE EPH FAMILY.  
DR ENBL: Z19110; G312902; -.  
DR PFAM: PF00041; fn3; 2.  
DR PFAM: PF00059; pkinase; 1.  
DR PFAM: PF00536; SAM; 1.  
DR PFAM: PF00536; SAM; 1.  
KW RECEPTOR; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;  
KW TRANSFERASE; GLYCOPROTEIN.  
FT NON\_TER 1 1  
FT NP\_BIND 593 601 ATP (BY SIMILARITY).  
FT ACT\_SITE 712 712 BY SIMILARITY.  
FT BINDING 619 619  
FT MOD\_RES 746 746 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 448 448 POTENTIAL.  
SQ SEQUENCE 952 AA; 105994 MW; D939D7D0 CRC32;

Query Match 31.2%; Score 82; DB 13; Length 952;  
Best Local Similarity 47.6%; Pred. No. 4.38e-02;  
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 289 LEWHPRETGGDDVYNYVC 309  
QY 7 IQWPPRPPEINKDDVQHKQVC 27

RESULT 8  
ID Q91571 PRELIMINARY; PRT; 985 AA.  
AC Q91571;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DE "Expression of an amphibian homolog of the Eph family of receptor  
tyrosine kinases is developmentally regulated.";  
DR ONCOGENE 10:1111-1117(1995).  
DR ENBL: U14164; G957215; -.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PFAM: PF00041; fn3; 2.  
DR PFAM: PF00069; pkinase; 1.  
DR PFAM: PF00536; SAM; 1.  
SQ SEQUENCE 985 AA; 110104 MW; 2416A8F9 CRC32;

Query Match 31.2%; Score 82; DB 13; Length 985;  
Best Local Similarity 42.9%; Pred. No. 4.38e-02;  
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 342 LEWHPRETGGDDVYNYVC 362  
QY 7 IQWPPRPPEINKDDVQHKQVC 27

RESULT 9  
ID Q91736 PRELIMINARY; PRT; 902 AA.  
AC Q91736;

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE EPH RECEPTOR TYROSINE KINASE (FRAGMENT).  
OS XENOPUS LAEVIS (AFRICAN CLAWED FROG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;  
OC MESOBATRACHIA; PIPOIDEA; PIPIDAE; XENOPODINAE; XENOPUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 96068901.  
RA SCALES J.B., WINNING R.S., RENAUD C.S., SHEA L.J., SARGENT T.D.;  
RT "Novel members of the eph receptor tyrosine kinase subfamily  
expressed during Xenopus development.";  
RL ONCOGENE 11:1745-1752(1995).  
DR ENBL: L43621; G974712; -.  
DR PROSITE: PS00790; RECEPTOR\_TYR\_KIN\_V\_1; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00791; RECEPTOR\_TYR\_KIN\_V\_2; 1.  
DR PFAM: PF00041; fn3; 2.  
DR PFAM: PF00069; pkinase; 1.  
DR PFAM: PF00536; SAM; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 902 AA; 100850 MW; C096D006 CRC32;

Query Match 30.0%; Score 79; DB 13; Length 902;  
Best Local Similarity 38.1%; Pred. No. 1.42e-01;  
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 258 LEWHPRETGGDDVYNYIC 278  
QY 7 IQWPPRPPEINKDDVQHKQVC 27

RESULT 10  
ID Q43569 PRELIMINARY; PRT; 943 AA.  
AC Q43569;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE EPH-LIKE RECEPTOR TYROSINE KINASE HEPHBLIC  
DE (EPH-LIKE RECEPTOR TYROSINE KINASE HEPHBLD).  
GN EPHBL  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY.  
RA STEIN E., SCHOCKLMANN H.O., DANIEL T.O.;  
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR ENBL: AF037333; G2739208; -.  
DR ENBL: AF037334; G2739210; -.  
SQ SEQUENCE 943 AA; 105291 MW; C4C78A21 CRC32;

Query Match 29.3%; Score 77; DB 4; Length 943;  
Best Local Similarity 38.1%; Pred. No. 3.05e-01;  
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Db 340 LEWHPRETGGDDVYNYIC 360  
QY 7 IQWPPRPPEINKDDVQHKQVC 27

RESULT 11  
ID Q06825 PRELIMINARY; PRT; 271 AA.  
AC Q06825;  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 28.7 KD PROTEIN.  
GN MTCY493.21C.  
OS MYCOBACTERIUM TUBERCULOSIS.

OC BACTERIA: FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;  
 OC ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA OLIVER K., HARRIS D.;  
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;  
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE: 96181548.  
 RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,  
 RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,  
 RA COLE S.T.;  
 RT "An integrated map of the genome of the tubercle bacillus,  
 RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium  
 RT leprae";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).  
 DR EMBL: Z95844; E318872; -.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 271 AA; D49C5921 CRC32;  
 Query Match 28.5%; Score 75; DB 2; Length 271;  
 Best Local Similarity 53.8%; Pred. No. 6.51e-01;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Db 90 WLDNDVVQWVDPDR 102  
 : : : : :  
 : : : : :  
 QY 1 WIENDEIQWNPDR 13  
 : : : : :  
 : : : : :  
 RESULT 12  
 ID Q05787 PRELIMINARY; PRT; 833 AA.  
 AC Q05787;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE CHROMOSOME XII COSMID 8167.  
 GN L8167.5.  
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEAE; SACCHAROMYCES.  
 [1]  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C (AB972);  
 RX MEDLINE: 97313267.  
 RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,  
 RA BENES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOFT A.,  
 RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEHLING O., HEUMANN K.,  
 RA HEUSS-NEITZEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,  
 RA LOUIS E.J., MESSENGUY F., MEWES H.W., MIOGA T., MOSTL D.,  
 RA MULLER-AUER S., NENTWICH U., OBERMAIER B., PIRAVANDI E., POHL T.M.,  
 RA PORTELLE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,  
 RA SCHARFE M., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,  
 RA UNDERWOOD A.P., URRESTARAZU L.A., VANDENBOL M., VERHASSELT P.,  
 RA VIERENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMBUUT R., WEDLER E.,  
 RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOEISEL J.D.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT XII";  
 RL NATURE 387:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C (AB972);  
 RA PAULY A.;  
 RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C (AB972);

RA WATERSTON R.;  
 RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C (AB972);  
 RA CHERRY J.M.;  
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: U14913; G544502; -.  
 SQ SEQUENCE 833 AA; 95480 MW; 5D133935 CRC32;  
 Query Match 27.8%; Score 73; DB 3; Length 833;  
 Best Local Similarity 33.3%; Pred. No. 1.37e+00;  
 Matches 8; Conservative 6; Mismatches 9; Indels 1; Gaps 1;  
 Db 684 WITREKVYWKPPSSPLNPNDTQH 707  
 : : : : :  
 : : : : :  
 QY 1 WIENDEIQWNPDRPEINKDDVOH 23  
 : : : : :  
 : : : : :  
 RESULT 13  
 ID Q23821 PRELIMINARY; PRT; 1368 AA.  
 AC Q23821;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE RECEPTOR TYROSINE KINASE.  
 OS CAENORHABDITIS VULGARIS.  
 OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
 OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 96177760.  
 RA SAKAI T., KOGA M., OHSHIMA Y.;  
 RT "Genomic structure and 5' regulatory regions of the let-23 gene in  
 RT the nematode C. elegans";  
 RL J. MOL. BIOL. 256:548-555(1996).  
 DR EMBL: D63427; D1010376; -.  
 DR PFAM: PF00069; pkinase; 1.  
 DR PFAM: PF00757; Furin-like; 1.  
 DR PFAM: PF01030; Recep\_Ldomain; 2.  
 SQ SEQUENCE 1368 AA; 156391 MW; 9BE7B0A0 CRC32;  
 Query Match 27.4%; Score 72; DB 5; Length 1368;  
 Best Local Similarity 36.7%; Pred. No. 1.98e+00;  
 Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 2;  
 Db 120 WIEANEIQWRNSTNQTVDDADIDLYKTVMF 149  
 : : : : :  
 : : : : :  
 QY 1 WIENDEIQWNPDRPEINKDDVOH-KQVCF 28  
 : : : : :  
 : : : : :  
 RESULT 14  
 ID Q35657 PRELIMINARY; PRT; 409 AA.  
 AC Q35657;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE G9 SIALIDASE PRECURSOR.  
 GN G9.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
 OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/BJ;  
 RX MEDLINE: 98029867.  
 RA CARRILLO M.B., MILNER C.M., BALL S.T., SNOEK M., CAMPBELL R.D.;  
 RT "Cloning and characterization of a sialidase from the murine  
 RT histocompatibility-2 complex: low levels of mRNA and a single amino  
 RT acid mutation are responsible for reduced sialidase activity in mice  
 RT carrying the Neula allele";  
 RL GLYCOBIOLOGY 7:975-986(1997).  
 DR EMBL: Y11412; E307743; -.

```

KW SIGNAL.
FT SIGNAL 1 39 POTENTIAL.
SQ SEQUENCE 409 AA; 44591 MW; FFB19BA5 CRC32;

Query Match 27.0%; Score 71; DB 11; Length 409;
Best Local Similarity 53.8%; Pred.No. 2.85e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 167 WSKDDGISWSPR 179
| : | : | : |
QY 1 WIENDEIQWNP 13

RESULT 15
ID 052973 PRELIMINARY; PRT: 1503 AA.
AC 052973;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CYCLOINULO-OLIGOSACCHARIDE FRUCTANOTRANSFERASE PRECURSOR.
OS BACILLUS CIRCULANS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC12554;
RA YAMADA K., MITSUI K., KUSHIBE S.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC12554;
RX MEDLINE: 98069491.
RA KANAI T., UEKI N., KAWAGUCHI T., TERANISHI Y., ATOMI H.,
RA TOMORBAATAR C., UEDA M., TANAKA A.;
RT "Recombinant thermostable cyclolinulo-oligosaccharide
RL fructanotransferase produced by Saccharomyces cerevisiae.";
RL APPL. ENVIRON. MICROBIOL. 63:4956-4960(1997).
DR EMBL: D87672; D1025272; -.
KW SIGNAL; TRANSFERASE.
FT SIGNAL 1 32 POTENTIAL.
SQ SEQUENCE 1503 AA; 167139 MW; D0397029 CRC32;

Query Match 27.0%; Score 71; DB 2; Length 1503;
Best Local Similarity 26.3%; Pred.No. 2.85e+00;
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 866 WVSDDMVHWNVRPALAPE 884
| : | : | : |
QY 1 WIENDEIQWNP 19

Search completed: Mon Aug 23 13:18:19 1999
Job time : 18 secs.

```

**This Page Blank (uspto)**

\*\*\*\*\*  
WQESNLA  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:08:40 1999; MasPar time 9.70 Seconds  
Tabular output not generated.

Title: >US-09-049-696-48  
Description: (1-35) from US09049696.pep  
Sequence: 1 WIENDEIQWNPPEINKDDVQHKVCFSTSSGG 35

Scoring table:  
TABLE unitprotatable  
Gap 60

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 2.032; Variance 0.600; scale 3.388

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description Pred. No.
1	35	100.0	228 20 W06548 Human colon specific 1.58e-37
2	35	100.0	228 29 W46879 Protein sequence enco 1.58e-37
3	6	17.1	877 14 R2657 Human mGLUR3. 2.98e-01
4	6	17.1	879 12 R4252 Human mGLUR3. 2.99e+01
5	6	17.1	2115 32 W59276 Rubella virus RA27/3 2.99e+01
6	6	17.1	2205 15 R79048 Infectious rubella vi 2.99e+01
7	5	14.3	29 3 R15584 Immunopeptide #3 deri 3.25e+02
8	5	14.3	30 3 R15585 Immunopeptide #4 deri 3.25e+02
9	5	14.3	48 3 R14752 Seroreactive epitope 3.25e+02
10	5	14.3	48 3 R4921 HCV-N (6957-7100) enc 3.25e+02
11	5	14.3	58 37 W74958 Human secreted protei 3.25e+02
12	5	14.3	59 38 W88625 Secreted protein enco 3.25e+02
13	5	14.3	91 29 W55449 H. pylori ORF 02ae116 3.25e+02
14	5	14.3	95 31 W27984 Staphylococcus aureus 3.25e+02
15	5	14.3	99 1 P81777 Sequence encoded by o 3.25e+02
16	5	14.3	103 29 W55209 H. pylori ORF 02ae116 3.25e+02

Protein sequence enco	3.25e+02
Sequence of tat prote	3.25e+02
TAT gene protein.	3.25e+02
Ehrlichia sp. E74.1 p	3.25e+02
Human secreted protei	3.25e+02
Amino acid sequence o	3.25e+02
Human secreted protei	3.25e+02
Homo sapiens SOCS11 p	3.25e+02
NANBH virus antigenic	3.25e+02
Tomato expansin leEX1	3.25e+02
S. pneumoniae transcr	3.25e+02
Arabidopsis expansin.	3.25e+02
Human GAP-43.	3.25e+02
Human neuronal growth	3.25e+02
Sequence encoded by h	3.25e+02
Human GAP-43.	3.25e+02
VP2 sequence for HRV	3.25e+02
Sequence of rhinoviru	3.25e+02
HPV 18 E7 protein fra	3.25e+02
Photosynthetic organ	3.25e+02
Rat cell growth regul	3.25e+02
Egr-1 transcription a	3.25e+02
Human secreted protei	3.25e+02
S. flexneri lcsA prot	3.25e+02
Arabidopsis thaliana	3.25e+02
Human cell growth reg	3.25e+02
Expandase mutant R74M	3.25e+02
Expandase mutant R74M	3.25e+02
Expandase mutant R266	3.25e+02
Expandase mutant R266	3.25e+02
Expandase mutant R74I	3.25e+02
Expandase mutant C155	3.25e+02
S. clavuligerus expan	3.25e+02
Expandase mutant R266	3.25e+02
S. clavuligerus expan	3.25e+02
Expandase mutant C155	3.25e+02
S. clavuligerus expan	3.25e+02
Expandase mutant R74F	3.25e+02
Expandase mutant C155	3.25e+02
S. clavuligerus expan	3.25e+02
S. clavuligerus expan	3.25e+02
Deacetoxyccephalospori	3.25e+02
S. clavuligerus wild	3.25e+02
S. clavuligerus expan	3.25e+02
Expandase mutant P157	3.25e+02
Expandase mutant P157	3.25e+02
Cephalosporin antibio	3.25e+02
Fas ligand (FasL) pro	3.25e+02
Murine Fas antigen ex	3.25e+02
MFAS sequence.	3.25e+02
Human fibrosarcoma ce	3.25e+02
Murine Fas.	3.25e+02
Human GRBP protein.	3.25e+02
Deacetoxyccephalospori	3.25e+02
Pseudomonas glutamina	3.25e+02
Human par-4 protein.	3.25e+02
D-alanine-D-alanine l	3.25e+02
Acetobacter xylinum C	3.25e+02
12D3 antigen sequence	3.25e+02
H. pylori cytoplasmic	3.25e+02
NANBH virus antigenic	3.25e+02
Human epidermoid carc	3.25e+02
NANB hepatitis virus	3.25e+02
Pectin-lyase-I.	3.25e+02
Ehrlichia sp. HGE-9 p	3.25e+02
Bacillus agaradherens	3.25e+02
Bacillus agaradherens	3.25e+02
Glycoprotein 50 (gp50)	3.25e+02
Pseudorabies virus gp	3.25e+02
Vpr protein binding H	3.25e+02
vpr binding protein H	3.25e+02

90	5	14.3	409 1	P81843	Sequence of alkaline	3.25e+02	163	953 18	R90923	F. bidentis pyruvate	3.25e+02
91	5	14.3	410 21	W12378	P300-CelB fusion cons	3.25e+02	164	966 23	W19916	Drosophila melanogast	3.25e+02
92	5	14.3	411 21	W12379	P300-CelB fusion cons	3.25e+02	165	985 21	W06786	ILTV glycoprotein g60	3.25e+02
93	5	14.3	411 21	W12381	P300-CelB fusion cons	3.25e+02	166	985 19	W00634	ILTV glycoprotein g60	3.25e+02
94	5	14.3	412 21	W12380	P300-CelB fusion cons	3.25e+02	167	995 4	R22234	Sequence encoded by I	3.25e+02
95	5	14.3	421 10	R55205	Bovine zona pellucida	3.25e+02	168	999 26	W26683	Bovine lysosomal alph	3.25e+02
96	5	14.3	421 37	W81815	Bovine ZPC protein.	3.25e+02	169	999 26	W26682	Bovine lysosomal alph	3.25e+02
97	5	14.3	434 18	R23215	CORK potassium channe	3.25e+02	170	999 26	W26684	Bovine lysosomal alph	3.25e+02
98	5	14.3	441 17	R94538	Drosophila Src28C tyr	3.25e+02	171	1081 30	W33817	Arabidopsis cellulose	3.25e+02
99	5	14.3	449 36	W30559	Aspergillus oryzae he	3.25e+02	172	1081 30	W33820	Arabidopsis cellulose	3.25e+02
100	5	14.3	454 38	W85046	Gamma subunit of a DN	3.25e+02	173	1098 19	R11139	B lymphocyte membrane	3.25e+02
101	5	14.3	458 18	R99316	Human SH-PTP1 variant	3.25e+02	174	1146 39	R93932	Haemophilus adhesin	3.25e+02
102	5	14.3	458 14	R72506	Porcine vasoactive in	3.25e+02	175	1243 12	R67708	Histidine kinase Cgsl	3.25e+02
103	5	14.3	462 11	R58592	Hepatitis C virus 800	3.25e+02	176	1243 12	R67708	Insulin receptor subs	3.25e+02
104	5	14.3	462 32	W57433	Cloned alkaline endog	3.25e+02	177	1297 30	W40200	Infected cell protein	3.25e+02
105	5	14.3	464 38	W85045	Gamma subunit of a DN	3.25e+02	178	1297 30	W40200	The amino acid sequen	3.25e+02
106	5	14.3	466 13	R71910	Erythroid p55.	3.25e+02	179	1404 7	R83804	Sequence of a serrate	3.25e+02
107	5	14.3	467 19	W00382	Bacillus cellulase BC	3.25e+02	180	1612 24	W24094	Ras-binding protein p	3.25e+02
108	5	14.3	467 20	W05731	Cellulase.	3.25e+02	181	1612 12	R66457	Chimeric ALL-1/AF-6 p	3.25e+02
109	5	14.3	479 3	P50034	Sequence encoded by t	3.25e+02	182	1810 16	R94563	Chicken cytotactin.	3.25e+02
110	5	14.3	479 1	R81013	Complete sequence of	3.25e+02	183	2039 32	W56322	Haemophilus paragalli	3.25e+02
111	5	14.3	496 37	W76984	Mouse Egr-1 protein.	3.25e+02	184	2042 32	W56319	Haemophilus paragalli	3.25e+02
112	5	14.3	509 1	P91060	Predicted amino acid	3.25e+02	185	2164 2	P81045	Sequence of the viral	3.25e+02
113	5	14.3	511 39	W73506	VSV-G protein sequenc	3.25e+02	186	2164 2	P81031	Peptides translated f	3.25e+02
114	5	14.3	511 19	W04539	Vesiculovirus glycopr	3.25e+02	187	24305 2	R24305	Translation of ORF 1	3.25e+02
115	5	14.3	513 18	R99314	Rat M1PTP (protein ty	3.25e+02	188	2458 2	R40311	Full length T4 encode	3.25e+02
116	5	14.3	521 5	R27808	Bovine herpes virus t	3.25e+02	189	2458 2	R07640	Deduced protein sequ	3.25e+02
117	5	14.3	521 8	R41344	Bovine herpesvirus ty	3.25e+02	190	2842 12	R63508	Adenomatous polyposi	3.25e+02
118	5	14.3	521 14	R77400	BHV1 gIII glycoprotei	3.25e+02	191	2843 35	W76144	Human APC protein #2.	3.25e+02
119	5	14.3	529 38	W85044	Tau subunit of a DNA	3.25e+02	192	2843 35	W76140	Human APC protein #1.	3.25e+02
120	5	14.3	533 1	P93113	Egr-1.	3.25e+02	193	2843 12	R58634	Adenomatous polyposi	3.25e+02
121	5	14.3	533 11	R63129	Mouse Egr-1 clone OC3	3.25e+02	194	2843 26	W35392	Human adenomatous Pol	3.25e+02
122	5	14.3	547 13	R67630	Non-A Non-B hepatitis	3.25e+02	195	2843 27	W38370	Human adenomatous Pol	3.25e+02
123	5	14.3	555 20	W10167	Nematode regulator of	3.25e+02	196	2843 21	W11922	Adenomatous polyposi	3.25e+02
124	5	14.3	573 38	W68490	Ehrlichia sp. E82.1 p	3.25e+02	197	2843 5	R26052	APC gene product in f	3.25e+02
125	5	14.3	594 34	W68092	Mouse neuronal PAS do	3.25e+02	198	2860 12	R63507	Adenomatous polyposi	3.25e+02
126	5	14.3	593 18	R99312	Human SH-PTP1 (protei	3.25e+02	199	2973 37	W76821	Human APC protein.	3.25e+02
127	5	14.3	619 5	R25672	Rat dopamine transpor	3.25e+02	200	3033 6	R33538	HNABH virus strain HC	3.25e+02
128	5	14.3	619 7	R36395	Dopamine transporter	3.25e+02	201	7 26	W34364	PKB substrate #11.	2.89e+03
129	5	14.3	620 8	R44492	Predicted sequence of	3.25e+02	202	8 21	W21080	Lipolytic enzyme opt.	2.89e+03
130	5	14.3	629 30	W38816	Arabidopsis cellulose	3.25e+02	203	9 14	R83817	SH3-binding proline-r	2.89e+03
131	5	14.3	634 30	W44193	Trichoderma harzianum	3.25e+02	204	9 28	W46384	Peptide sequence of t	2.89e+03
132	5	14.3	634 32	W60210	Trichoderma harzianum	3.25e+02	205	9 20	W01695	HCV NS3 domain protea	2.89e+03
133	5	14.3	635 14	R75940	Human myeloproliferat	3.25e+02	206	11 26	W25467	SH3 domain binding re	2.89e+03
134	5	14.3	635 4	R23970	MPLV env protein with	3.25e+02	207	12 21	W21076	Lipolytic enzyme opt.	2.89e+03
135	5	14.3	635 4	R22352	Bacterial alpha-1,3-g	3.25e+02	208	12 24	W23199	Terminal peptide exte	2.89e+03
136	5	14.3	636 29	W41509	Aspergillus oryzae 5-	3.25e+02	209	12 21	R92713	Mink p27 Kipl tryptic	2.89e+03
137	5	14.3	636 36	W30557	Aspergillus oryzae 5-	3.25e+02	210	12 21	W11137	HIV-1 Tat protein pep	2.89e+03
138	5	14.3	636 29	W41498	5-aminolevulinic acid	3.25e+02	211	13 39	W81643	Mouse elf-1 N-termina	2.89e+03
139	5	14.3	649 38	W81977	Ehrlichia sp. B3 prot	3.25e+02	212	13 39	W86105	Peptide from humanise	2.89e+03
140	5	14.3	736 25	W26653	Maize male fertility	3.25e+02	213	13 39	W86092	Peptide from humanise	2.89e+03
141	5	14.3	748 39	W89272	Granulocytic Ehrlich	3.25e+02	214	13 39	W67994	Fragment of human sec	2.89e+03
142	5	14.3	780 30	W37051	S. putrefaciens EPO b	3.25e+02	215	13 39	W86100	Peptide from humanise	2.89e+03
143	5	14.3	780 20	R99463	Biosynthetic enzyme o	3.25e+02	216	14 15	R93511	Class II +5 GST-LYN S	2.89e+03
144	5	14.3	780 9	R42454	Enzyme involved in ei	3.25e+02	217	14 15	R93515	Class II +5 GST-YES S	2.89e+03
145	5	14.3	785 1	R04907	Type I interferon rec	3.25e+02	218	14 15	R93534	Class II +5 GST-PI3K	2.89e+03
146	5	14.3	835 17	R96206	Invasin protein.	3.25e+02	219	15 15	R93466	GST-PI3K protein tyro	2.89e+03
147	5	14.3	843 38	W85030	Green fluorescent pro	3.25e+02	220	15 26	W38905	Peptide resembling an	2.89e+03
148	5	14.3	848 33	W62600	Starch branching enzy	3.25e+02	221	15 26	W38905	Peptide resembling an	2.89e+03
149	5	14.3	849 11	R63070	Human EAA3d excitator	3.25e+02	222	16 32	W54430	Human P512 protein s	2.89e+03
150	5	14.3	853 38	W75021	Protein tyrosine phos	3.25e+02	223	16 27	W25386	PLC gamma SH3 domain	2.89e+03
151	5	14.3	856 37	W75231	Human secreted protei	3.25e+02	224	16 27	W37665	pppmy motif containin	2.89e+03
152	5	14.3	861 32	W43074	HIV-1 gp120 protein f	3.25e+02	225	17 39	W67994	Fragment of human sec	2.89e+03
153	5	14.3	865 11	R63069	Human EAA3c excitator	3.25e+02	226	17 14	R76149	hML(152-168).	2.89e+03
154	5	14.3	905 11	R63068	Human EAA3b excitator	3.25e+02	227	18 14	R83829	Proline-rich peptide	2.89e+03
155	5	14.3	905 16	R75883	Human EAA3a receptor (	3.25e+02	228	18 14	R76150	hML(152-169).	2.89e+03
156	5	14.3	905 11	R60112	Human EAA3a excitator	3.25e+02	229	20 10	R50300	Anti-bacterial glycop	2.89e+03
157	5	14.3	920 2	R11993	Glutamate receptor 5-	3.25e+02	230	21 29	W46691	Biotinylation peptide	2.89e+03
158	5	14.3	921 20	W06077	Drosophila rATA-bindi	3.25e+02	231	28 18	R88861	Peptide targetted to	2.89e+03
159	5	14.3	921 11	P56487	TATA-binding protein-	3.25e+02	232	28 32	W48702	Bovine growth-inducin	2.89e+03
160	5	14.3	921 23	W25028	TATA-binding protein	3.25e+02	233	29 26	R67646	Chymotrypsin-based ca	2.89e+03
161	5	14.3	951 13	R75704	Eph-related CER6.	3.25e+02	234	30 17	R85014	Peptide rII from the	2.89e+03
162	5	14.3	953 14	R76707	Recombinant cold-resi	3.25e+02	235	30 17	R85006	Peptide rII from the	2.89e+03



382	4	11.4	236 3	R15690	V1-Lab-Vh constructio	2.89e+03	2.89e+03	281 36	W72190	HSV-2 strain SB5 Cont	2.89e+03
383	4	11.4	236 3	R14702	Vh-Lab-V1 constructio	2.89e+03	2.89e+03	281 2	P70294	Subtype adr HBSAg P31	2.89e+03
384	4	11.4	236 31	W28089	Amino acid sequence o	2.89e+03	2.89e+03	281 2	W85102	Thyroid hormone recep	2.89e+03
385	4	11.4	237 39	W95440	A33/212 single-chain	2.89e+03	2.89e+03	288 38	W89984	Expressed antigen for	2.89e+03
386	4	11.4	237 20	W09238	HCV solubilised NS3 p	2.89e+03	2.89e+03	289 38	W85092	Thyroid hormone recep	2.89e+03
387	4	11.4	238 33	W61950	Green fluorescent pro	2.89e+03	2.89e+03	293 9	R52669	Equine herpesvirus gl	2.89e+03
388	4	11.4	238 25	W24232	Aequorea victoria gre	2.89e+03	2.89e+03	295 20	W11867	Mutant tissue factor,	2.89e+03
389	4	11.4	238 35	W76371	A. victoria green flu	2.89e+03	2.89e+03	296 39	W87887	Protein encoded by th	2.89e+03
390	4	11.4	238 26	W22059	Aequorea victoria pro	2.89e+03	2.89e+03	296 39	W89198	Aspartate-specific cy	2.89e+03
391	4	11.4	238 20	W05307	Green fluorescent pro	2.89e+03	2.89e+03	299 13	W46386	Protease sequence of	2.89e+03
392	4	11.4	238 30	W52327	GRF mutant f66W/1123V	2.89e+03	2.89e+03	299 13	R68549	Hepatitis C virus (HC	2.89e+03
393	4	11.4	238 30	W52334	Engineered green fluo	2.89e+03	2.89e+03	300 38	W82558	C. elegans glycoengin	2.89e+03
394	4	11.4	238 35	W59822	Green fluorescent pro	2.89e+03	2.89e+03	300 39	W73518	KSHV assembly protein	2.89e+03
395	4	11.4	238 35	W59821	Green fluorescent pro	2.89e+03	2.89e+03	300 22	W18643	Fragmented human amyl	2.89e+03
396	4	11.4	238 33	W65078	A. victoria green flu	2.89e+03	2.89e+03	301 23	W21697	Maize proRIP.	2.89e+03
397	4	11.4	238 35	W59812	Green fluorescent pro	2.89e+03	2.89e+03	302 30	W50200	A krill derived prote	2.89e+03
398	4	11.4	238 32	W42485	A. victoria green flu	2.89e+03	2.89e+03	303 38	W89842	Protein encoded by cl	2.89e+03
399	4	11.4	238 32	W42501	A. victoria green flu	2.89e+03	2.89e+03	303 9	R52667	Equine herpesvirus US	2.89e+03
400	4	11.4	238 32	W42494	A. victoria green flu	2.89e+03	2.89e+03	304 38	W89919	Antigen 4 from cluste	2.89e+03
401	4	11.4	238 32	W21751	Mutant Aequorea victor	2.89e+03	2.89e+03	305 39	W73519	KSHV assembly protein	2.89e+03
402	4	11.4	238 32	W42488	A. victoria green flu	2.89e+03	2.89e+03	307 6	R29843	HCV NS2-NS4 peptide N	2.89e+03
403	4	11.4	238 32	W42496	A. victoria green flu	2.89e+03	2.89e+03	307 4	R22248	Sequence of rye-grass	2.89e+03
404	4	11.4	238 32	W42486	A. victoria green flu	2.89e+03	2.89e+03	309 38	W89960	Antigen from cluster	2.89e+03
405	4	11.4	238 30	W05209	Green fluorescent pro	2.89e+03	2.89e+03	309 39	W83322	Single chain Apo-2 an	2.89e+03
406	4	11.4	239 32	W42502	A. victoria green flu	2.89e+03	2.89e+03	310 39	W83324	Single chain Apo-2 an	2.89e+03
407	4	11.4	241 32	W42504	A. victoria green flu	2.89e+03	2.89e+03	310 32	W60610	Human mucosal adressi	2.89e+03
408	4	11.4	241 33	W95441	A33/218 single-chain	2.89e+03	2.89e+03	312 39	W83323	Single chain Apo-2 an	2.89e+03
409	4	11.4	241 33	W65077	A. victoria green flu	2.89e+03	2.89e+03	312 2	P60471	Portion of a human T-	2.89e+03
410	4	11.4	241 32	W42483	A. victoria green flu	2.89e+03	2.89e+03	313 2	R10690	Cephalosporin antibio	2.89e+03
411	4	11.4	242 33	W59809	Green fluorescent pro	2.89e+03	2.89e+03	313 32	W60611	Human mucosal adressi	2.89e+03
412	4	11.4	243 39	W86002	Murine anti-5T4 anti	2.89e+03	2.89e+03	314 27	W27286	Rabbit enteropathogen	2.89e+03
413	4	11.4	243 3	R14598	V1-Lab-Vh constructio	2.89e+03	2.89e+03	315 4	R22214	Sequence of interleuk	2.89e+03
414	4	11.4	243 34	W63152	27 kDa Tuberculosis m	2.89e+03	2.89e+03	318 38	W89997	Expressed antigen for	2.89e+03
415	4	11.4	244 19	R97774	Human thrombopoietin	2.89e+03	2.89e+03	319 37	W80598	Human neurotensin rec	2.89e+03
416	4	11.4	246 16	R79873	Anti-Egfr single chai	2.89e+03	2.89e+03	319 2	P71199	Sequence of Hrbv-III	2.89e+03
417	4	11.4	246 8	R40924	Protein able to bind	2.89e+03	2.89e+03	320 36	W77349	Human telomere repeat	2.89e+03
418	4	11.4	247 38	W89037	Polypeptide fragment	2.89e+03	2.89e+03	321 2	R06434	Spl-like protein enco	2.89e+03
419	4	11.4	247 6	R32091	H and L chain Fv regi	2.89e+03	2.89e+03	323 19	W00988	Human interleukin-1 b	2.89e+03
420	4	11.4	248 12	R66999	Trypsin-like protease	2.89e+03	2.89e+03	323 33	W62042	Rat arginase.	2.89e+03
421	4	11.4	248 15	R88470	Pre-pro-trypsin.	2.89e+03	2.89e+03	323 9	R45268	ICE mutant E324X.	2.89e+03
422	4	11.4	249 34	W71029	Acid phosphatase enzy	2.89e+03	2.89e+03	325 29	W47420	Micrococcus luteus pr	2.89e+03
423	4	11.4	250 21	W01648	HCV NS3 soluble prote	2.89e+03	2.89e+03	326 38	R81983	Ehrlichia sp. E46.1 p	2.89e+03
424	4	11.4	251 6	R28067	Sequence encoded by t	2.89e+03	2.89e+03	326 11	R58304	Rat ciliary neurotroph	2.89e+03
425	4	11.4	253 38	W81586	CTLA4/E7 fusion prote	2.89e+03	2.89e+03	326 39	W89279	Granulocytic Ehrlich	2.89e+03
426	4	11.4	253 38	W87562	CTLA4-E7 fusion prote	2.89e+03	2.89e+03	329 1	R05552	Murine interferon reg	2.89e+03
427	4	11.4	253 4	R22323	Marek Disease Virus U	2.89e+03	2.89e+03	329 38	W85076	Esterase E101 from Th	2.89e+03
428	4	11.4	254 33	W64209	Oleosin-metallothione	2.89e+03	2.89e+03	330 38	W86020	Human homeobox gene H	2.89e+03
429	4	11.4	256 38	W81985	Ehrlichia sp. E80.3 p	2.89e+03	2.89e+03	332 39	W84217	Thrombopoietin protei	2.89e+03
430	4	11.4	256 4	R22584	ScFvB18 construct #3.	2.89e+03	2.89e+03	332 20	W12950	Mpl ligand analogue,	2.89e+03
431	4	11.4	256 39	W89780	Staphylococcus aureus	2.89e+03	2.89e+03	332 20	W12921	Mpl ligand analogue,	2.89e+03
432	4	11.4	256 4	R22582	ScFvB18 construct mut	2.89e+03	2.89e+03	332 4	R22213	Sequence of interleuk	2.89e+03
433	4	11.4	256 4	R22583	ScFvB18 construct #2.	2.89e+03	2.89e+03	333 39	W84353	Partial human ubiquit	2.89e+03
434	4	11.4	256 13	R67914	H. pylori haemaggluti	2.89e+03	2.89e+03	333 25	W33107	Chicken lactic acid d	2.89e+03
435	4	11.4	257 38	W88578	Secreted protein enco	2.89e+03	2.89e+03	334 38	W89915	Antigen 2 from cluste	2.89e+03
436	4	11.4	257 9	R46897	GGF2BPP3.	2.89e+03	2.89e+03	334 38	W82627	Ehrlichia sp. HGE-15	2.89e+03
437	4	11.4	257 25	W35085	E. coli serine acetyl	2.89e+03	2.89e+03	337 28	R35287	Human G52-24 secret	2.89e+03
438	4	11.4	258 3	P60008	Sequence of oncogene	2.89e+03	2.89e+03	337 7	W25228	VD44 region of human	2.89e+03
439	4	11.4	258 23	W08487	C6 human sfv antibody	2.89e+03	2.89e+03	340 17	R85863	WD-40 domain-contg. h	2.89e+03
440	4	11.4	260 39	W89101	Anti-digoxin single c	2.89e+03	2.89e+03	343 21	W01561	MOC1 protein.	2.89e+03
441	4	11.4	262 18	R97726	B10 single chain T-ce	2.89e+03	2.89e+03	344 38	W83139	Br20 protein.	2.89e+03
442	4	11.4	263 21	W02306	Aequorea victoria Gre	2.89e+03	2.89e+03	346 10	R60642	psts variant.	2.89e+03
443	4	11.4	264 21	W12344	Human insulin-like gr	2.89e+03	2.89e+03	346 10	R60644	psts variant.	2.89e+03
444	4	11.4	265 38	W89987	Expressed antigen for	2.89e+03	2.89e+03	346 10	R60653	psts variant.	2.89e+03
445	4	11.4	266 33	W55069	Streptococcus pneumon	2.89e+03	2.89e+03	346 10	R60654	psts variant.	2.89e+03
446	4	11.4	267 39	W67885	Human secreted protei	2.89e+03	2.89e+03	347 10	R54865	Thermus thermophilus	2.89e+03
447	4	11.4	270 38	W88784	Polypeptide fragment	2.89e+03	2.89e+03	347 14	R61986	Porphyromonas gengiva	2.89e+03
448	4	11.4	270 39	W95369	Rat interleukin (IL)-	2.89e+03	2.89e+03	347 14	R61988	Porphyromonas gengiva	2.89e+03
449	4	11.4	270 39	W85366	Murine interleukin (IL)	2.89e+03	2.89e+03	351 10	R51687	MVP-5180/91 fragmen	2.89e+03
450	4	11.4	276 38	W88934	Polypeptide fragment	2.89e+03	2.89e+03	352 39	W73786	M. tuberculosis anti	2.89e+03
451	4	11.4	276 23	W21711	RIP fusion protein. R	2.89e+03	2.89e+03	352 39	W73676	M. tuberculosis anti	2.89e+03
452	4	11.4	278 1	R05162	Sequence of human bon	2.89e+03	2.89e+03	353 15	R81356	Thrombopoietin protei	2.89e+03
453	4	11.4	278 20	W06081	Drosophila TATA-bindi	2.89e+03	2.89e+03	353 15	R79907	Haematopoietic protei	2.89e+03
454	4	11.4	280 17	R87455	BPP2 glial growth fac	2.89e+03	2.89e+03	354 36	W72122	HSV-2 strain SB5 Cont	2.89e+03



528	4	11.4	354 35	W72045	HSV-2 strain SB5 Cont	2.89e+03	601	4	11.4	414 39	W86143	Streptokinase (SK) pr	2.89e+03
529	4	11.4	355 23	W10100	Human C-C chemokine r	2.89e+03	602	4	11.4	415 34	W69228	Human lysosomal sial	2.89e+03
530	4	11.4	355 32	W51746	Human C-C chemokine r	2.89e+03	603	4	11.4	417 29	W44330	Class II acyl-ACP thi	2.89e+03
531	4	11.4	355 32	W51746	Human C-C chemokine r	2.89e+03	604	4	11.4	418 14	R77364	T. longibrachiatum en	2.89e+03
532	4	11.4	356 26	W25943	Human CCR3 chemokine	2.89e+03	605	4	11.4	419 38	W83397	Maize DnaJ clone prot	2.89e+03
533	4	11.4	358 38	W89971	Antigen from chloster	2.89e+03	606	4	11.4	419 19	W01037	Mycoplasma 46-48 kba	2.89e+03
534	4	11.4	358 38	W89815	Protein encoded by cl	2.89e+03	607	4	11.4	420 33	W55075	Streptococcus pneumo	2.89e+03
535	4	11.4	360 38	W70804	Amino acid sequence o	2.89e+03	608	4	11.4	422 38	W73363	Chicken PAF-AH protei	2.89e+03
536	4	11.4	360 36	W72212	HSV-2 strain SB5 Cont	2.89e+03	609	4	11.4	422 26	W23004	Canine herpesvirus un	2.89e+03
537	4	11.4	362 3	R12463	HLA-Bw53 exon.	2.89e+03	610	4	11.4	423 39	W73657	M. tuberculosis antiq	2.89e+03
538	4	11.4	362 25	W16642	Truncated HSV specifi	2.89e+03	611	4	11.4	423 39	W73657	M. tuberculosis antiq	2.89e+03
539	4	11.4	364 38	W81984	Ehrlichia sp. E46.2 p	2.89e+03	612	4	11.4	423 39	W73657	M. tuberculosis antiq	2.89e+03
540	4	11.4	364 39	W89280	Granulocytic Ehrlich	2.89e+03	613	4	11.4	423 17	R91305	Transcription factor-	2.89e+03
541	4	11.4	364 39	W53356	IL-1 alpha propeptide-E	2.89e+03	614	4	11.4	424 28	W43038	Amino acid sequence o	2.89e+03
542	4	11.4	365 35	W69697	Human coxsackievirus	2.89e+03	615	4	11.4	425 21	W14763	Mouse soluble kit lig	2.89e+03
543	4	11.4	365 31	W57212	Human coxsackievirus	2.89e+03	616	4	11.4	427 1	P90478	Polyptide of human	2.89e+03
544	4	11.4	365 26	W23536	CDK inhibitory fusion	2.89e+03	617	4	11.4	428 19	R99388	Human REST protein (a	2.89e+03
545	4	11.4	367 38	W67472	Human retroviral rece	2.89e+03	618	4	11.4	431 37	W80310	Human PACAP/VIP R-2B	2.89e+03
546	4	11.4	367 36	W80549	Rat methadone-specifi	2.89e+03	619	4	11.4	433 39	W89784	Staphylococcus aureus	2.89e+03
547	4	11.4	367 13	R87671	Mouse opioide receptor	2.89e+03	620	4	11.4	434 38	W85053	dhah gene product of	2.89e+03
548	4	11.4	367 8	R40115	APP-HCV-E2 fusion pro	2.89e+03	621	4	11.4	434 26	R03743	Monkey transforming g	2.89e+03
549	4	11.4	367 13	R74298	Mouse kappa-3 oploid	2.89e+03	622	4	11.4	435 3	R13907	Human protein phospho	2.89e+03
550	4	11.4	371 38	W67443	C. jejuni flagellin/E	2.89e+03	623	4	11.4	435 11	R56483	ScFV PRAS108 and PRAS	2.89e+03
551	4	11.4	372 39	W67857	Human secreted protei	2.89e+03	624	4	11.4	437 39	W87771	UDP-N-acetylmuramate:	2.89e+03
552	4	11.4	372 30	W52251	Rat galanin receptor	2.89e+03	625	4	11.4	437 36	W77350	Altered telomere repe	2.89e+03
553	4	11.4	374 39	W88207	Alcohol dehydrogenase	2.89e+03	626	4	11.4	438 39	W83392	Caenorhabditis elegan	2.89e+03
554	4	11.4	374 39	W88201	Alcohol dehydrogenase	2.89e+03	627	4	11.4	439 38	W82645	Ehrlichia sp. extende	2.89e+03
555	4	11.4	374 39	W88200	Alcohol dehydrogenase	2.89e+03	628	4	11.4	439 36	W77348	Human telomere repeat	2.89e+03
556	4	11.4	374 39	W88202	Alcohol dehydrogenase	2.89e+03	629	4	11.4	439 37	W59995	Cell division Divlb p	2.89e+03
557	4	11.4	374 39	W88208	Horse liver alcohol d	2.89e+03	630	4	11.4	440 39	W86320	Kidney injury associa	2.89e+03
558	4	11.4	374 39	W88209	Horse liver alcohol d	2.89e+03	631	4	11.4	440 17	R93118	Human GABA-A receptor	2.89e+03
559	4	11.4	374 39	W88205	Alcohol dehydrogenase	2.89e+03	632	4	11.4	441 36	W72744	Drosophila kinesin N-	2.89e+03
560	4	11.4	374 39	W88206	Alcohol dehydrogenase	2.89e+03	633	4	11.4	442 38	W85035	Green fluorescent pro	2.89e+03
561	4	11.4	374 39	W88203	Alcohol dehydrogenase	2.89e+03	634	4	11.4	442 17	W00505	N.meningitidis IM2394	2.89e+03
562	4	11.4	374 39	W88204	Alcohol dehydrogenase	2.89e+03	635	4	11.4	443 39	W84264	Protein encoded by cp	2.89e+03
563	4	11.4	375 38	W40085	Human membrane fusion	2.89e+03	636	4	11.4	443 38	W89850	Antigen 1 from cluste	2.89e+03
564	4	11.4	375 31	W46287	Rhizobium meliloti DN	2.89e+03	637	4	11.4	443 26	W24025	Single chain antigen	2.89e+03
565	4	11.4	376 24	W26404	Human fibromodulin.	2.89e+03	638	4	11.4	445 14	R72451	Brevibacterium flavum	2.89e+03
566	4	11.4	378 36	W69428	Human secreted protei	2.89e+03	639	4	11.4	445 23	W06588	B. lactofermentum arg	2.89e+03
567	4	11.4	378 27	W35864	Human FAS-ligand:IG2	2.89e+03	640	4	11.4	446 39	W95506	Mortierella alpina de	2.89e+03
568	4	11.4	379 38	W85036	Alpha-actinin actin-b	2.89e+03	641	4	11.4	446 38	W85119	A delta-5 desaturase	2.89e+03
569	4	11.4	379 27	W33195	Corn barnacle G-prote	2.89e+03	642	4	11.4	447 5	R24779	Rabbit Gnt 1.	2.89e+03
570	4	11.4	380 39	W89183	S. pneumoniae GlmU OR	2.89e+03	643	4	11.4	448 10	R51479	Human TLE-4.	2.89e+03
571	4	11.4	380 32	P93309	Promoter of beta-isop	2.89e+03	644	4	11.4	451 39	W89977	Protein encoded by cl	2.89e+03
572	4	11.4	382 39	W88491	Human liver clone HP0	2.89e+03	645	4	11.4	451 16	R88904	Human insulin recepto	2.89e+03
573	4	11.4	385 19	W02682	G-protein coupled rat	2.89e+03	646	4	11.4	454 33	W56773	Homo sapiens PSP1-1 c	2.89e+03
574	4	11.4	386 7	R75592	Sequence of delta-15	2.89e+03	647	4	11.4	458 37	W70517	Candida tropicalis IP	2.89e+03
575	4	11.4	388 1	P50389	N-terminal human seru	2.89e+03	648	4	11.4	459 39	W89182	S. pneumoniae GlmU po	2.89e+03
576	4	11.4	389 24	W25151	Modified Arthrobacter	2.89e+03	649	4	11.4	461 19	R98341	Human p57 protein.	2.89e+03
577	4	11.4	389 24	W25151	Modified Arthrobacter	2.89e+03	650	4	11.4	462 3	R12401	Enantioselective amid	2.89e+03
578	4	11.4	389 36	W71461	A modified sarcosine	2.89e+03	651	4	11.4	463 3	R14179	Human serum albumin l	2.89e+03
579	4	11.4	389 1	P90330	N-terminal human seru	2.89e+03	652	4	11.4	464 14	R65166	Biliary glycoprotein.	2.89e+03
580	4	11.4	390 20	W07397	Bacillus sarcosine ox	2.89e+03	653	4	11.4	466 38	W82398	Mouse G3BP protein.	2.89e+03
581	4	11.4	390 1	P50331	N-terminal human seru	2.89e+03	654	4	11.4	466 4	P40032	Gamma anti-carcinoemb	2.89e+03
582	4	11.4	392 21	W01763	1-153 delta 112-115 c	2.89e+03	655	4	11.4	469 39	W89560	Triticum sp. cysteine	2.89e+03
583	4	11.4	393 24	W32867	Mitogen activated pro	2.89e+03	656	4	11.4	469 31	W56793	L. lactis Fl portion	2.89e+03
584	4	11.4	393 30	W37149	Murine Ena-Vasp like	2.89e+03	657	4	11.4	469 20	R15510	Tomato ACC synthase e	2.89e+03
585	4	11.4	394 27	W34541	S-adenosylmethionine	2.89e+03	658	4	11.4	469 3	W04722	Aromatic acyl transfe	2.89e+03
586	4	11.4	395 20	W09038	Mitogen-activated pro	2.89e+03	659	4	11.4	470 13	R77847	ADP-glucose-pyrophosp	2.89e+03
587	4	11.4	395 33	W57571	Homo sapiens BAP152 s	2.89e+03	660	4	11.4	470 13	R66240	ADPGPP large subunit.	2.89e+03
588	4	11.4	396 21	W01764	1-153 c-mpl ligand-mo	2.89e+03	661	4	11.4	471 25	W16794	Aspartase-470.	2.89e+03
589	4	11.4	397 39	W82731	Adenovirus pCSCG2SAR	2.89e+03	662	4	11.4	472 39	W89559	Triticum sp. cysteine	2.89e+03
590	4	11.4	401 24	W15225	Abl interactor protei	2.89e+03	663	4	11.4	474 3	P60181	Sequence of ARV-2 env	2.89e+03
591	4	11.4	402 7	R34439	Sequence of glycoprot	2.89e+03	664	4	11.4	475 8	R39750	Delta endotoxin.	2.89e+03
592	4	11.4	403 39	W87639	A hypersensitive resp	2.89e+03	665	4	11.4	479 24	W11891	MACH isoform alpha.	2.89e+03
593	4	11.4	403 34	W62455	Kidney injury associa	2.89e+03	666	4	11.4	481 9	R47234	Rubella E1 glycoprote	2.89e+03
594	4	11.4	405 39	W86329	Snake venom fibrolase	2.89e+03	667	4	11.4	483 3	R12522	Factor Xa-cleavable h	2.89e+03
595	4	11.4	405 1	P50458	Actinomadura hibiscla	2.89e+03	668	4	11.4	483 3	R12885	Factor Xa-cleavable h	2.89e+03
596	4	11.4	407 31	W54381	pHCV423-encoded APP-H	2.89e+03	669	4	11.4	487 9	R49829	Highly heat-resistant	2.89e+03
597	4	11.4	410 14	R79227	Drosophila kinesin N-	2.89e+03	670	4	11.4	487 4	R22380	Antigen mc-35c.	2.89e+03
598	4	11.4	411 36	W72745	Snake venom fibrolase	2.89e+03	671	4	11.4	488 39	R66004	Human B7-1.5t4.1 prot	2.89e+03
599	4	11.4	411 1	P90457	Snake venom fibrolase	2.89e+03	672	4	11.4	491 7	R36979	Histamine H1 receptor	2.89e+03
600	4	11.4	414 39	W86144	De-immunised streptok	2.89e+03	673	4	11.4	491 33	W62316	Mouse protein phospho	2.89e+03
										492 21	W01560	CYB1 protein.	2.89e+03

674	4	11.4	433 39	W84068	Human mitofusin Fzo h	2.89e+03	747	590 38	W82649	Ehrlichia sp. extende	2.89e+03
675	4	11.4	494 28	W42996	Putative mature potas	2.89e+03	748	592 39	W82661	A. thaliana L-galacto	2.89e+03
676	4	11.4	495 33	W49026	Saccharomyces cerevis	2.89e+03	749	592 38	W70797	Human interleukin-6R-	2.89e+03
677	4	11.4	496 20	W01411	Human TIE-2 ligand 2.	2.89e+03	750	594 38	W89552	Secreted protein enco	2.89e+03
678	4	11.4	496 3	R13713	PKV glycoprotein gA.	2.89e+03	751	594 36	W72094	HSV-2 strain S85 Cont	2.89e+03
679	4	11.4	498 3	R13916	Pseudorabies virus gX	2.89e+03	752	595 39	W86003	Anti-574 single chain	2.89e+03
680	4	11.4	498 19	W06440	Glyceral kinase.	2.89e+03	753	595 38	W85021	mPKAc-green fluoresce	2.89e+03
681	4	11.4	499 32	W47529	Amino acid sequence o	2.89e+03	754	596 39	W89184	Osf2/Cbfa1 native pol	2.89e+03
682	4	11.4	500 4	R21682	HSV-1 vP16- derived f	2.89e+03	755	596 36	W77099	P. chrysogenum beta-N	2.89e+03
683	4	11.4	501 33	W44868	TGF-beta superfamily	2.89e+03	756	599 39	W95462	Anti-Tac(Fv)-PE40 fus	2.89e+03
684	4	11.4	501 23	W31880	GFP variants p4-3 and	2.89e+03	757	604 38	W85016	Erk2-green fluorescen	2.89e+03
685	4	11.4	501 15	R69600	New TGF-beta family m	2.89e+03	758	605 38	W85007	Erk2-green fluorescen	2.89e+03
686	4	11.4	502 38	W81585	CTLA4/p97 fusion prot	2.89e+03	759	605 34	W62838	Glycine max antimicro	2.89e+03
687	4	11.4	502 38	W87561	CTLA4-p97 fusion prot	2.89e+03	760	606 38	W85019	p38-green fluorescen	2.89e+03
688	4	11.4	502 23	W21720	RIP fusion protein, R	2.89e+03	761	607 38	W85010	p38-green fluorescen	2.89e+03
689	4	11.4	502 33	W69216	V274T variant human a	2.89e+03	762	609 39	W85187	Vibriolysin, usefuf f	2.89e+03
690	4	11.4	503 33	W56302	Babesia microti MN-10	2.89e+03	763	609 37	W83215	Human h-NUMB-R.	2.89e+03
691	4	11.4	505 14	R73963	Humicola insolens pro	2.89e+03	764	610 8	R39510	Chimeric human serum	2.89e+03
692	4	11.4	507 9	R47517	MEL EPO receptor.	2.89e+03	765	610 29	W46733	Amino acid sequence o	2.89e+03
693	4	11.4	508 31	W51254	Arabidopsis proto-por	2.89e+03	766	613 37	W73009	Cobra venom protease	2.89e+03
694	4	11.4	509 30	W40480	Rat SH2 binding prote	2.89e+03	767	616 38	W83023	Activator of leukotox	2.89e+03
695	4	11.4	509 17	R97646	Human SOX-9.	2.89e+03	768	618 23	W19583	Human apoptosis inhib	2.89e+03
696	4	11.4	509 20	W11977	Hiv sf2 gp120.	2.89e+03	769	618 23	W19746	Human inhibitor of ap	2.89e+03
697	4	11.4	513 15	R79945	Helicobacter pylori a	2.89e+03	770	619 39	W73532	MEK2 protein.	2.89e+03
698	4	11.4	514 38	W81976	Ehrlichia sp. W20.2 p	2.89e+03	771	619 30	W56162	A mitogen-activated p	2.89e+03
699	4	11.4	515 33	W55073	Streptococcus pneumon	2.89e+03	772	620 36	W72179	HSV-2 strain S85 Cont	2.89e+03
700	4	11.4	516 39	W67617	P. chrysogenum phenyl	2.89e+03	773	621 27	W36451	Human tyrosine phosph	2.89e+03
701	4	11.4	521 33	W62317	Mouse protein phosph	2.89e+03	774	622 23	W11543	Human preprothrombin	2.89e+03
702	4	11.4	521 13	R68973	ADP-glucose-pyrophosp	2.89e+03	775	623 38	W88707	Secreted protein enco	2.89e+03
703	4	11.4	521 4	R24529	Enantioselective amid	2.89e+03	776	623 16	R86875	Cancer suppressor tra	2.89e+03
704	4	11.4	521 13	R69997	HIV-1 env protein.	2.89e+03	777	624 38	W85015	Erk1-green fluorescen	2.89e+03
705	4	11.4	522 1	P82676	GAGROD sequence from	2.89e+03	778	625 23	W18008	Arabidopsis CER1 prot	2.89e+03
706	4	11.4	522 26	W34638	Human occludin.	2.89e+03	779	626 29	W41501	Human DP.75, a putati	2.89e+03
707	4	11.4	523 36	W22776	Human septin-2 protei	2.89e+03	780	629 38	W67473	Human retroviral rece	2.89e+03
708	4	11.4	524 39	W95521	Protease activated pr	2.89e+03	781	630 38	W85018	Jnk1-green fluorescen	2.89e+03
709	4	11.4	526 38	W83138	FL-CEA protein.	2.89e+03	782	631 38	W85006	Erk1-green fluorescen	2.89e+03
710	4	11.4	527 30	W48419	Amino acid sequence o	2.89e+03	783	633 38	W85009	Jnk1-green fluorescen	2.89e+03
711	4	11.4	528 34	W69963	Aeromonas caviae xyla	2.89e+03	784	634 29	W34537	Nudaurelia beta virus	2.89e+03
712	4	11.4	529 38	W80498	A protein designated	2.89e+03	785	635 38	W85034	Green fluorescent pro	2.89e+03
713	4	11.4	529 39	W73473	Human cancer-related	2.89e+03	786	638 5	R03924	E. coli HSP (dnak).	2.89e+03
714	4	11.4	529 38	W85055	Human protein designa	2.89e+03	787	640 32	W48855	Luciferase-lysostaphi	2.89e+03
715	4	11.4	530 38	W82656	Ehrlichia sp. extende	2.89e+03	788	642 36	W71645	Fluorescent calmoduli	2.89e+03
716	4	11.4	533 29	W55476	H. pylori ORF 07ap202	2.89e+03	789	642 34	W62596	Human nuclear protein	2.89e+03
717	4	11.4	534 8	R39571	Sequence of G-FIT.	2.89e+03	790	642 36	W71647	Fluorescent calmoduli	2.89e+03
718	4	11.4	535 20	W11255	Hippocampus calcium/c	2.89e+03	791	647 38	W73376	Human HPDDV78 protein	2.89e+03
719	4	11.4	539 31	W57336	WBp1 protein fragment	2.89e+03	792	649 39	W67883	Human secreted protei	2.89e+03
720	4	11.4	540 24	W00978	Phosphatidylinositol-	2.89e+03	793	652 36	W71646	Fluorescent calmoduli	2.89e+03
721	4	11.4	541 39	W86006	Candida glabrata hist	2.89e+03	794	655 30	W48265	Sf9 alpha-mannosidase	2.89e+03
722	4	11.4	544 38	W85028	CDK2-green fluorescen	2.89e+03	795	658 17	R85856	WD-40 domain-contg. C	2.89e+03
723	4	11.4	544 38	W85029	Green fluorescent pro	2.89e+03	796	666 37	W72911	Mycobacterium tubercu	2.89e+03
724	4	11.4	545 39	W89277	Granulocytic Ehrlichi	2.89e+03	797	669 5	R28620	Lys256 deleted acetoh	2.89e+03
725	4	11.4	545 36	W72253	HSV-2 strain S85 Cont	2.89e+03	798	669 5	R28622	Asp376 deleted acetoh	2.89e+03
726	4	11.4	545 8	R41726	Peptide derived from	2.89e+03	799	670 31	W59147	A. thaliana acetolact	2.89e+03
727	4	11.4	546 38	W81975	Ehrlichia sp. W20.1 p	2.89e+03	800	671 20	W08988	A. thaliana acetolact	2.89e+03
728	4	11.4	546 32	W56523	Helicobacter pylori g	2.89e+03	801	672 15	R73536	MEK1 protein.	2.89e+03
729	4	11.4	548 5	R05788	Luciferase.	2.89e+03	802	672 35	W77544	MEK1 protein.	2.89e+03
730	4	11.4	548 27	W27377	L. cruciata luciferas	2.89e+03	803	672 19	R77544	Mammalian MEK kinase	2.89e+03
731	4	11.4	548 3	R14119	Luciferase.	2.89e+03	804	672 12	R66029	Human growth arrest s	2.89e+03
732	4	11.4	549 12	R66030	Mammalian MEK kinase	2.89e+03	805	678 28	W48463	Human GFAT.	2.89e+03
733	4	11.4	554 34	W69223	HAP4 protein.	2.89e+03	806	681 8	R43348	Human GFAT.	2.89e+03
734	4	11.4	555 13	R77864	S. clavuligerus ORF7	2.89e+03	807	685 37	W08813	Nucleotide sequence o	2.89e+03
735	4	11.4	557 39	R81996	A. thaliana trehalase	2.89e+03	808	686 3	R14538	Hepatitis C Virus pro	2.89e+03
736	4	11.4	559 12	R64764	E. coli groEL protein	2.89e+03	809	686 39	W86162	Insect GABA receptor	2.89e+03
737	4	11.4	560 12	R71294	Human glycoprotein V.	2.89e+03	810	695 11	R58923	Mouse amyloid precurs	2.89e+03
738	4	11.4	567 7	R36601	TGF-beta1 receptor ty	2.89e+03	811	702 38	W83137	CEA protein.	2.89e+03
739	4	11.4	573 36	W76317	Myceliophthora thermo	2.89e+03	812	708 20	W06411	Maltose binding prote	2.89e+03
740	4	11.4	573 33	W51782	Myceliophthora thermo	2.89e+03	813	708 33	W69844	Amino acid sequence o	2.89e+03
741	4	11.4	576 13	R72318	Laccase RSLac1.	2.89e+03	814	709 30	W53345	Human adrenilycitra	2.89e+03
742	4	11.4	591 39	W81993	35S Camv anti-sense t	2.89e+03	815	709 27	W35315	M. catarrhalis O8 tra	2.89e+03
743	4	11.4	591 11	R58859	Newcastle Disease Vir	2.89e+03	816	714 9	R47244	Fragment of 101 kd pr	2.89e+03
744	4	11.4	595 11	R58589	Hepatitis C virus 340	2.89e+03	817	716 6	R30730	B. burgdorferi 79 kd	2.89e+03
745	4	11.4	595 1	P90388	Mature human serum al	2.89e+03	818	718 38	W85024	Erk2-green fluorescen	2.89e+03
746	4	11.4	595 27	W31271	Human frizzled-5 prot	2.89e+03	819	718 17	R85865	WD-40 domain-contg. D	2.89e+03

820	4	11.4	719 38	W85012	Smad2-green flouresce	2.89e+03	893	4	11.4	917 30	W44836	Rat ICAM-4 protein.	2.89e+03
821	4	11.4	721 36	W49862	Thermotoga maritima M	2.89e+03	894	4	11.4	920 28	W25716	Mouse beta meltrin pr	2.89e+03
822	4	11.4	724 27	W36797	Novel human gene, des	2.89e+03	895	4	11.4	924 38	W73511	Human neuronal ICAM-4	2.89e+03
823	4	11.4	726 38	W85022	PKB-green human gene,	2.89e+03	896	4	11.4	924 32	W60161	Human intercellular a	2.89e+03
824	4	11.4	728 38	W85041	Green fluorescent pro	2.89e+03	897	4	11.4	930 32	W55961	Human transient recep	2.89e+03
825	4	11.4	738 15	W69852	Ethylene response (ET	2.89e+03	898	4	11.4	932 38	W85040	PKG-green fluorescent	2.89e+03
826	4	11.4	738 36	W73117	A. thaliana ethylene	2.89e+03	899	4	11.4	933 38	W85039	Green fluorescent pro	2.89e+03
827	4	11.4	740 9	W47174	Predicted sequence of	2.89e+03	900	4	11.4	933 39	W89200	Human phosphatidylcho	2.89e+03
828	4	11.4	743 10	W51477	Human TLE-2.	2.89e+03	901	4	11.4	936 39	W89801	Staphylococcus aureus	2.89e+03
829	4	11.4	748 36	W70457	Mutant human disinteg	2.89e+03	902	4	11.4	943 38	W83168	Human receptor tyrosi	2.89e+03
830	4	11.4	749 31	W56134	Mus musculus KUZ prot	2.89e+03	903	4	11.4	945 39	W73624	Human secreted protei	2.89e+03
831	4	11.4	749 17	W94895	Human trophobin.	2.89e+03	904	4	11.4	948 3	R14403	Extracellular domain	2.89e+03
832	4	11.4	749 9	W42485	Feline infectious per	2.89e+03	905	4	11.4	950 2	R06376	Product of the scl1 g	2.89e+03
833	4	11.4	751 21	W01825	Human metalloprotein	2.89e+03	906	4	11.4	961 39	W83367	Caenorhabditis eleg	2.89e+03
834	4	11.4	753 39	W83927	Human T85 protein.	2.89e+03	907	4	11.4	968 38	W85011	p85alpha-green floure	2.89e+03
835	4	11.4	754 18	R92149	HSA:Fc gamma RII fusi	2.89e+03	908	4	11.4	970 38	W85020	p85alpha-green floure	2.89e+03
836	4	11.4	763 30	W41734	Human TRAF-2 kinase.	2.89e+03	909	4	11.4	974 2	R06399	Lambda gt10ch201 enco	2.89e+03
837	4	11.4	765 38	W89914	Antigen from cluster	2.89e+03	910	4	11.4	976 30	W53571	Rat p138 protein.	2.89e+03
838	4	11.4	768 6	R29869	HCV NS2-NS4 peptide N	2.89e+03	911	4	11.4	986 31	W59050	H. contortus pHC-1 an	2.89e+03
839	4	11.4	770 23	W19491	APP70 mutant A-beta-	2.89e+03	912	4	11.4	990 39	W90075	P. vulgaris chondroit	2.89e+03
840	4	11.4	770 10	R51476	Human TLE-1.	2.89e+03	913	4	11.4	990 15	R77460	Chondroitinase-II.	2.89e+03
841	4	11.4	772 10	R51111	Human TLE-3.	2.89e+03	914	4	11.4	993 5	R27650	Human calcium channel	2.89e+03
842	4	11.4	772 10	R51478	Human TLE-3.	2.89e+03	915	4	11.4	996 38	W85033	IkappaB kinase alpha	2.89e+03
843	4	11.4	775 2	R06400	Lambda gt10ch2 encode	2.89e+03	916	4	11.4	997 38	W85032	Green fluorescent pro	2.89e+03
844	4	11.4	783 8	R39473	Prepro-HSA-G-CSF chim	2.89e+03	917	4	11.4	1004 32	W37956	Human polyhomeotic 1	2.89e+03
845	4	11.4	784 27	W34179	Human G binding prot	2.89e+03	918	4	11.4	1015 28	W34259	Rat ob receptor isofo	2.89e+03
846	4	11.4	790 18	R95565	N. meningitidis serot	2.89e+03	919	4	11.4	1020 19	W05141	scFv2(FRP5/225)-ETA (	2.89e+03
847	4	11.4	797 38	W85043	NFKappaB p65 subunit-	2.89e+03	920	4	11.4	1037 39	W67643	A serine/threonine pr	2.89e+03
848	4	11.4	797 38	W85042	PKB-green fluorescent	2.89e+03	921	4	11.4	1039 38	W85014	Stat5-green flouresce	2.89e+03
849	4	11.4	798 38	W85025	Smad4-green flouresce	2.89e+03	922	4	11.4	1040 39	W73584	Stat5-green flouresce	2.89e+03
850	4	11.4	805 36	W72248	HSV-2 strain SB5 Cont	2.89e+03	923	4	11.4	1041 20	W11866	RAG-1 protein.	2.89e+03
851	4	11.4	806 38	W85013	Smad4-green flouresce	2.89e+03	924	4	11.4	1045 38	W85050	ARH1 gene product, At	2.89e+03
852	4	11.4	813 30	W40059	Mouse P300/CBP-associ	2.89e+03	925	4	11.4	1057 14	R79653	Stat5-green flouresce	2.89e+03
853	4	11.4	816 13	R66931	AMML chromosome inv(1	2.89e+03	926	4	11.4	1061 38	W87504	Human E1 protein.	2.89e+03
854	4	11.4	821 36	W72250	HSV-2 strain SB5 Cont	2.89e+03	927	4	11.4	1063 14	R73916	Human N-methyl-D-aspa	2.89e+03
855	4	11.4	822 32	W69847	Amino acid sequence o	2.89e+03	928	4	11.4	1065 30	W33819	Rubella virus strain	2.89e+03
856	4	11.4	822 32	W56490	Zea mays starch blanc	2.89e+03	929	4	11.4	1088 39	W88253	Arabidopsis cellulose	2.89e+03
857	4	11.4	836 38	W85017	Grk5-green flourescen	2.89e+03	930	4	11.4	1088 39	W89564	Gracilariaopsis lemame	2.89e+03
858	4	11.4	842 39	W85008	Grk5-green flourescen	2.89e+03	931	4	11.4	1088 14	R72398	Human parathyroid cal	2.89e+03
859	4	11.4	847 39	W85603	Hexosaminidase enzyme	2.89e+03	932	4	11.4	1091 33	W63145	Glucan Lyase 1.	2.89e+03
860	4	11.4	849 3	R14925	Mutant SP6DNA polymer	2.89e+03	933	4	11.4	1092 39	W88257	Human calcium channel	2.89e+03
861	4	11.4	854 32	R43058	Hiv-1 gp120 protein f	2.89e+03	934	4	11.4	1092 15	R91234	Algal alpha-1,4-gluc	2.89e+03
862	4	11.4	854 32	W43070	Hiv-1 gp120 protein f	2.89e+03	935	4	11.4	1110 33	W63148	Human calcium channel	2.89e+03
863	4	11.4	855 32	W43069	Hiv-1 gp120 protein f	2.89e+03	936	4	11.4	1171 38	W85037	Human fluorescent pro	2.89e+03
864	4	11.4	856 8	R41027	Selectively deglycosy	2.89e+03	937	4	11.4	1174 38	W67438	Human protein tyrosin	2.89e+03
865	4	11.4	856 8	R41031	Selectively deglycosy	2.89e+03	938	4	11.4	1176 8	R39755	Delta endotoxin.	2.89e+03
866	4	11.4	856 8	R41029	Selectively deglycosy	2.89e+03	939	4	11.4	1179 2	R08258	B. thuringiensis toxin	2.89e+03
867	4	11.4	856 8	R41026	Selectively deglycosy	2.89e+03	940	4	11.4	1181 38	W85038	NFAT1-green flouresce	2.89e+03
868	4	11.4	858 30	W56018	Recombinant botulinum	2.89e+03	941	4	11.4	1181 18	R91307	Helicobacter pylori T	2.89e+03
869	4	11.4	859 3	R12893	OmpAL streptokinase-s	2.89e+03	942	4	11.4	1188 6	R29870	HCV NS2-NS4 peptide M	2.89e+03
870	4	11.4	859 13	R72594	H. pylori tagA antige	2.89e+03	943	4	11.4	1195 39	W75420	Human sapiens cdo tumo	2.89e+03
871	4	11.4	862 11	R60563	Yeast 2.6 kb agglutin	2.89e+03	944	4	11.4	1205 16	R93930	T.thermophilus nitrat	2.89e+03
872	4	11.4	863 3	P60349	HTLV-III virus (HIV v	2.89e+03	945	4	11.4	1206 3	P60679	Bovine endothelial ni	2.89e+03
873	4	11.4	863 6	R28955	Non-cleavable, solubl	2.89e+03	946	4	11.4	1212 38	W87503	Sequence of Rift Vall	2.89e+03
874	4	11.4	865 38	W85037	Zap70-green flouresce	2.89e+03	947	4	11.4	1220 21	W13882	Human N-methyl-D-aspa	2.89e+03
875	4	11.4	865 2	R70175	Sequence encoded by e	2.89e+03	948	4	11.4	1225 30	W52289	8603c toxin.	2.89e+03
876	4	11.4	870 11	R59924	Human GAP protein.	2.89e+03	949	4	11.4	1229 39	W87633	Human sapiens cdo tumo	2.89e+03
877	4	11.4	871 38	W85026	Green fluorescent pro	2.89e+03	950	4	11.4	1247 39	W73535	CryET5 protein sequen	2.89e+03
878	4	11.4	873 14	R79459	Infectious Laryngotra	2.89e+03	951	4	11.4	1268 30	W58774	MEKK5 protein.	2.89e+03
879	4	11.4	880 38	W73352	HIV envelope protein.	2.89e+03	952	4	11.4	1274 39	W89253	Human ALP.	2.89e+03
880	4	11.4	880 27	W23333	Human immunodeficienc	2.89e+03	953	4	11.4	1288 24	W26328	Human breast cancer g	2.89e+03
881	4	11.4	880 13	R77846	Mouse Rse rPRK.	2.89e+03	954	4	11.4	1289 6	R28810	Mouse alpha-1 collage	2.89e+03
882	4	11.4	886 39	W89787	Staphylococcus aureus	2.89e+03	955	4	11.4	1291 21	W13885	Mouse alpha-1 collage	2.89e+03
883	4	11.4	888 25	W19857	Human semaphorin 2.	2.89e+03	956	4	11.4	1291 35	W68392	17b toxin.	2.89e+03
884	4	11.4	890 39	W73534	MEKK4 protein.	2.89e+03	957	4	11.4	1310 11	R60392	Clostridium botulinum	2.89e+03
885	4	11.4	890 13	R77845	Human Rse rPRK.	2.89e+03	958	4	11.4	1310 11	R60392	Varicella zoster viru	2.89e+03
886	4	11.4	897 39	W73474	Human retinoblastoma	2.89e+03	959	4	11.4	1336 38	W87511	Human N-methyl-D-aspa	2.89e+03
887	4	11.4	900 26	W22105	Murine leptin recepto	2.89e+03	960	4	11.4	1342 3	R13833	HER-3 epithelial grow	2.89e+03
888	4	11.4	903 28	W25715	Mouse alpha meltrin p	2.89e+03	961	4	11.4	1354 34	W71020	A modified Rho target	2.89e+03
889	4	11.4	909 10	R50092	Humanised anti-CEA sF	2.89e+03	962	4	11.4	1367 11	R63122	IGF-1 receptor.	2.89e+03
890	4	11.4	913 31	W53829	Pseudomonas OrfX secr	2.89e+03	963	4	11.4	1367 17	R91429	Human type I insulin-	2.89e+03
891	4	11.4	916 38	W85023	Amino acid sequence o	2.89e+03	964	4	11.4	1376 17	R93601	Kaposi's sarcoma asso	2.89e+03
892	4	11.4	917 38	W73512	Rat ICAM-4 protein se	2.89e+03	965	4	11.4	1384 36	W72224	HSV-2 strain SB5 Cont	2.89e+03
										1385 6	R28889	Toxin 17a.	2.89e+03

Matches	35;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Db	45	wiendeiqwnpprpeinkddvqhkvcfartssgg	79						
QY	1	wiendeiqwnpprpeinkddvqhkvcfartssgg	35						

RESULT 2  
 ID W46879 standard; Protein: 228 AA.  
 AC W46879;  
 DT 22-JUN-1998 (first entry)  
 DE Protein sequence encoded by a colon-specific gene.  
 KW Colon-specific gene; probe; detection; expression; human;  
 KW diagnostic assay; Colon cancer; antibody; screening.  
 OS Homo sapiens.  
 PN US5733748-A.  
 PD 31-MAR-1998.  
 PF 06-JUN-1995; 459667.  
 PR 06-JUN-1995; US-469667.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen C, Yu G;  
 DR WPI: 98-229823/20.  
 DR N-PSDB; V16572.  
 PT Colon-specific nucleic acids - useful as probes for detecting colon  
 cancer micrometastases  
 PT Claim 1; Fig 5; 51pp; English.  
 PS W46876-80 and W4682-85 represent proteins encoded by colon-specific  
 genes. The polynucleotides encoding these proteins can be used  
 as probes to detect expression of the corresponding human genes,  
 e.g. in diagnostic assays for detecting micrometastases of colon cancer.  
 CC Recombinant cells containing the polynucleotides can be used to  
 produce the proteins, in order that antibodies can be raised and  
 used in further screening or diagnostics.  
 CC sequence 228 AA;  
 SQ

## ALIGNMENTS

## ALIGNMENTS

Db	45	wiendeiqwnpprpeinkddvqhkvcfstsgg	79
QY	1	wiendeiqwnpprpeinkddvqhkvcfstsgg	35

RESULT		3	
ID	R82657	standard; protein: 877 AA.	
AC	R82657;		
DT	20-DEC-1995	(first entry)	
DE	Human mglur3.		
KW	Metabotropic glutamate receptor 3; mglur3; stroke; epilepsy;		
KW	Alzheimer disease; detection; diagnosis; therapy.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	domain	575..597	
FT		/label= TMD-I	
FT		/note= "transmembrane domain I"	
FT	domain	612..632	
FT		/label= TMD-II	
FT		/note= "transmembrane domain II"	
FT	domain	644..662	
FT		/label= TMD-III	
FT		/note= "transmembrane domain III"	
FT	domain	686..707	
FT		/label= TMD-IV	
FT		/note= "transmembrane domain IV"	
FT	domain	733..754	
FT		/label= TMD-V	
FT		/note= "transmembrane domain V"	
FT	domain	768..789	
FT		/label= TMD-VI	
FT		/note= "transmembrane domain VI"	
FT	domain	802..826	
FT		/label= TMD-VII	
FT			

```

FI  /note= "TRANSmembrane domain VI"
FT 802..826
FT  /label= TMD-VII

```

FT /note= "transmembrane domain VII"

PN WO9522609-A2.  
 PD 24-AUG-1995.  
 PF 21-FEB-1995; G00356.  
 PR 21-FEB-1994; GB-003285.  
 PR 01-AUG-1994; GB-015532.  
 PA (WELL ) WELLCOME FOUND LTD.  
 PI Makoff AJ;  
 DR WPI: 95-302715/39.  
 DR N-PSDB; T03887.  
 PT New isolated human metabotropic glutamate receptors - used for  
 PT detection, diagnosis and therapy of diseases associated with the  
 PT receptors, eg. stroke, epilepsy and Alzheimer's disease.  
 PS Claim 2; Page 32-35; 55pp; English.  
 CC mRNA from the human brain median frontal cortex and amygdala was used  
 CC to construct a cDNA library in lambda ZAP II. cDNA was amplified by  
 CC PCR primers (T03889-95) based on rat mglur1-5 sequences to obtain  
 CC cDNA encoding human mglur3.  
 CC Sequence 877 AA;  
 SQ

Query Match 17.1%; Score 6; DB 14; Length 877;  
 Best Local Similarity 100.0%; Pred. No. 2.99e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 einkdd 81  
 QY 15 EINKDD 20

RESULT 4  
 ID R64252 standard; Protein; 879 AA.  
 AC R64252;  
 DT 21-JUL-1995 (first entry)  
 DE Human mglur3.  
 KW Metabotropic glutamate receptor; mglur3; hippocampus; cerebellum;  
 KW PCMV-T7-3; PCMV-T7-2.  
 OS Homo sapiens.  
 PN WO9429449-A.  
 PD 22-DEC-1994.  
 PF 03-JUN-1994; U06273.  
 PR 04-JUN-1993; US-072574.  
 PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.  
 PI Daggett L, Ellis SB, Hess SD, Johnson EC, Liaw C;  
 PI Pontsler A;  
 DR WPI: 95-036478/05.  
 DR N-PSDB; Q80418.  
 PT New DNA encoding human metabotropic glutamate receptor  
 PT sub-type(s) - and related proteins, probes, RNA, transformed  
 PT cells and antibodies, useful in treatment, diagnosis and  
 PT identification of specific modulators  
 PS Claim 11; Page 83-87; 125pp; English.  
 CC Human hippocampus and cerebellum cDNA libraries were screened with  
 CC DNA encoding the rat mglur2 receptor. Several clones were isolated  
 CC and these were ligated to obtain a full-length construct (given in  
 CC Q80418) encoding human mglur3 (R64252). The mglur3 coding sequence  
 CC is operatively linked to the regulatory elements of PCMV-T7-3 or  
 CC PCMV-T3-2 for expression in mammalian cells.  
 SQ Sequence 879 AA;

Query Match 17.1%; Score 6; DB 12; Length 879;  
 Best Local Similarity 100.0%; Pred. No. 2.99e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 einkdd 83  
 QY 15 EINKDD 20

RESULT 5  
 ID W59276 standard; Protein; 2115 AA.  
 AC W59276;  
 DT 11-SEP-1998 (first entry)  
 DE Rubella virus RA27/3 NSP-ORF non-structural proteins.

KW genomic sequence; viral genome; rubella; RA27/3; vaccine; CRS;  
 KW congenital rubella syndrome; vaccination; nonstructural protein;  
 KW NS-ORF.  
 OS Rubella virus.  
 PN WO9820901-A1.  
 PD 22-MAY-1998.  
 PF 07-NOV-1997; U20399.  
 PR 12-NOV-1996; US-030734.  
 PA (UYGE-) UNIV GEORGIA STATE RES FOUND.  
 PI Frey TK, Pougatchev KV;  
 DR WPI: 98-297616/26.  
 DR N-PSDB; V34766.  
 PT Genomic sequence for RA27/3 strain of rubella virus - useful to  
 PT identify further attenuated strains of rubella virus, e.g. for use  
 PT in vaccines for congenital rubella syndrome  
 PS Disclosure; Fig 2; 34pp; English.  
 CC The sequence is that encoded by the NSP-ORF, it comprises non-  
 CC structural proteins that function primarily in viral RNA replication.  
 CC Sequence 2115 AA;  
 SQ

Query Match 17.1%; Score 6; DB 32; Length 2115;  
 Best Local Similarity 100.0%; Pred. No. 2.99e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 475 pprpei 480  
 QY 11 PPRPEI 16

RESULT 6  
 ID R79048 standard; Protein; 2205 AA.  
 AC R79048;  
 DT 27-FEB-1996 (first entry)  
 DE Infectious rubella virus N-terminal transcript.  
 KW Rubella; vaccine; mutant; epitope; virus; autoimmune disease;  
 KW pregnancy; foetal infection; vector; plasmid.  
 OS Rubella virus.  
 PN US5439814-A.  
 PD 08-AUG-1995.  
 PF 28-JUN-1991; US-722334.  
 PR 28-JUN-1991; US-722334.  
 PR 19-JUL-1993; US-093453.  
 PA (GEOR-) GEORGIA STATE RES FOUND INC.  
 PI Dominguez G, Frey TK, Wang C;  
 DR WPI: 95-283097/37.  
 DR N-PSDB; Q97686.  
 PT New DNA encoding infectious rubella virus - esp. non-pathogenic  
 PT mutant virus for use in vaccines having reduced side effects  
 PS Disclosure; Columns 17-28; 27pp; English.  
 CC Non-pathogenic mutants of the DNA corresponding to this sequence are  
 CC useful in vaccines (which may include epitopes from other viruses).  
 CC The mutant vaccines are less likely to cause foetal infections.  
 CC autoimmune disease or neurological symptoms, so can be administered  
 CC safely to older or pregnant women. The mutant sequence is  
 CC preferably present in a vector, especially a bacterial plasmid that  
 CC allows replication of the sequence.  
 SQ Sequence 2205 AA;

Query Match 17.1%; Score 6; DB 15; Length 2205;  
 Best Local Similarity 100.0%; Pred. No. 2.99e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 474 pprpei 479  
 QY 11 PPRPEI 16

RESULT 7  
 ID R15584 standard; Protein; 29 AA.  
 AC R15584;  
 DT 02-MAR-1992 (first entry)  
 DE Immunopeptide #3 derived from HPV18 E7 peptide.  
 KW cervical cancer; cervical intraepithelial neoplasia; CIN; wart;

KW squamous cell carcinoma; ELISA; HPV 18.  
OS Synthetic.  
PN WO9118294-A.  
PD 28-NOV-1991.  
PF 13-MAY-1991; SE0335.  
PR 11-MAY-1990; SE-001705.  
PA (MEDS-) MEDSCAND AB.  
PI Dillner J, Dillner L, Cheng HM;  
DR WPI: 91-369390/50.  
PT Diagnosis of human papilloma virus infection and PV-carrying  
PT tumours - using synthetic peptide(s) to detect virus specific  
PT antigen-antibody complexes by immunoassay  
PS Disclosure; Page 39; 72pp; English.  
CC This is one of a large number of peptides which have been  
CC synthesised on the basis of the amino acid sequences for the E2, E4,  
CC E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The  
CC selection of peptide sequences was based on the assumption that an  
CC immunoreactive region might be situated in the same relative region  
CC of a protein from different HPV types. The peptides were used in  
CC diagnostic immunoassays to detect HPV-infection.  
CC See R15523-R15601.  
CC Sequence 29 AA;  
SQ

Query Match 14.3%; Score 5; DB 3; Length 29;  
Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 20 endei 24  
|||||  
QY 3 ENDEI 7

RESULT 8  
ID R15585 standard; Protein; 30 AA.  
AC R15585;  
DT 02-MAR-1992 (first entry)  
DE Immunopeptide #4 derived from HPV18 E7 peptide.  
KW cervical cancer; cervical intraepithelial neoplasia; CIN; wart;  
KW squamous cell carcinoma; ELISA; HPV 18.  
OS Synthetic.  
PN WO9118294-A.  
PD 28-NOV-1991.  
PF 13-MAY-1991; SE0335.  
PR 11-MAY-1990; SE-001705.  
PA (MEDS-) MEDSCAND AB.  
PI Dillner J, Dillner L, Cheng HM;  
DR WPI: 91-369390/50.  
PT Diagnosis of human papilloma virus infection and PV-carrying  
PT tumours - using synthetic peptide(s) to detect virus specific  
PT antigen-antibody complexes by immunoassay  
PS Disclosure; Page 39; 72pp; English.  
CC This is one of a large number of peptides which have been  
CC synthesised on the basis of the amino acid sequences for the E2, E4,  
CC E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The  
CC selection of peptide sequences was based on the assumption that an  
CC immunoreactive region might be situated in the same relative region  
CC of a protein from different HPV types. The peptides were used in  
CC diagnostic immunoassays to detect HPV-infection.  
CC See R15523-R15601.  
CC Sequence 30 AA;  
SQ

Query Match 14.3%; Score 5; DB 3; Length 30;  
Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 6 endei 10  
|||||  
QY 3 ENDEI 7

DT 28-JAN-1992 (first entry)  
DE Seroreactive epitope #1 of HPV 18 protein E7.  
KW vaccine; HPV18.  
OS Synthetic.  
PN EP-456197-A.  
PD 13-NOV-1991.  
PF 07-MAY-1991; 107423.  
PR 10-MAY-1990; DE-015044.  
PA (BEHW) BEHRINGERWERKE AG.  
PI Bleul C, Gissmann L, Muller M;  
DR WPI: 91-334182/45.  
PT New sero-reactive epitope(s) of human papilloma virus 18 proteins  
PT - and corresp. proteins and antibodies, useful in vaccines and  
PT for diagnosis  
PS Claim 3; Page 5; 8pp; German.  
CC This peptide can be used to identify antibodies specific to E7  
CC protein. It is also useful for a vaccine against HPV18. The E7  
CC epitope was identified using anti-E7 serum to screen a recombinant  
CC phage expression library containing 100bp fragments of HPV18 DNA.  
CC The DNA inserts from positive phage were sequenced and based on  
CC this information, overlapping decapeptides were prepared on  
CC polyethylene pins. They were tested (ELISA) against the antisera  
CC to locate the epitopes. See R14748-R14753.  
CC Sequence 48 AA;  
SQ

Query Match 14.3%; Score 5; DB 3; Length 48;  
Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 endei 30  
|||||  
QY 3 ENDEI 7

RESULT 10  
ID R14921 standard; Protein; 48 AA.  
AC R14921;  
DT 03-JAN-1992 (first entry)  
DE HCV-N (6957-7100) encoded epitope.  
KW Epitope; PCR; diagnosis.  
OS Hepatitis C virus.  
PN WO9114779-A.  
PD 03-OCT-1991.  
PR 28-MAR-1991; J00405.  
PR 28-MAR-1990; JP-080185.  
PR 13-JUN-1990; JP-154230.  
PR 14-JUN-1990; JP-153979.  
PR 09-NOV-1990; JP-305795.  
PA (MITK) MITSUI TOATSU CHEM INC.  
PI Takada T, Enomoto N, Date T, Nakao T;  
DR WPI: 91-310579/42.  
DR N-PSDB: Q14771.  
PT New nucleotide sequences encoding HCV epitope(s) - for diagnosis  
PT of hepatitis C virus infection via polymerase chain reaction  
PS Disclosure; Fig 1(H); 79pp; Japanese.  
CC The nucleotide sequences represented in Q14076-86 and Q14767-71  
CC encode epitopes from structural, non-structural and 5' untranslated  
CC domains of hepatitis C virus. The sequences are used for accurate  
CC and simple diagnosis and typing of HCV infection, using PCR  
CC amplification techniques.  
CC Sequence 48 AA;  
SQ

Query Match 14.3%; Score 5; DB 3; Length 48;  
Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 pprpe 9  
|||||  
QY 11 PPRPE 15

RESULT 9  
ID R14752 standard; Protein; 48 AA.  
AC R14752;

AC W74958;  
 DT 23-JAN-1999 (first entry)  
 DE Human secreted protein encoded by gene 82 clone HPRAX55.  
 KW Human; secreted protein; testis; tumour; foetal brain tissue;  
 KW fusion protein; cancer; central nervous system; seizure;  
 KW diagnosis; neurodegenerative disease.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Misc-difference 5 /label= unknown  
 FT WO9839448-A2.  
 PN 11-SEP-1998: U04493.  
 PD 06-MAR-1998: U04493.  
 PF 02-OCT-1997: US-061060.  
 PR 07-MAR-1997: US-038621.  
 PR 07-MAR-1997: US-040161.  
 PR 07-MAR-1997: US-040162.  
 PR 07-MAR-1997: US-040163.  
 PR 07-MAR-1997: US-040333.  
 PR 07-MAR-1997: US-040334.  
 PR 07-MAR-1997: US-040336.  
 PR 07-MAR-1997: US-040626.  
 PR 11-APR-1997: US-043311.  
 PR 11-APR-1997: US-043312.  
 PR 11-APR-1997: US-043313.  
 PR 11-APR-1997: US-043314.  
 PR 11-APR-1997: US-043568.  
 PR 11-APR-1997: US-043569.  
 PR 11-APR-1997: US-043576.  
 PR 11-APR-1997: US-043578.  
 PR 11-APR-1997: US-043580.  
 PR 11-APR-1997: US-043669.  
 PR 11-APR-1997: US-043670.  
 PR 11-APR-1997: US-043671.  
 PR 11-APR-1997: US-043672.  
 PR 11-APR-1997: US-043674.  
 PR 23-MAY-1997: US-047492.  
 PR 23-MAY-1997: US-047500.  
 PR 23-MAY-1997: US-047501.  
 PR 23-MAY-1997: US-047502.  
 PR 23-MAY-1997: US-047503.  
 PR 23-MAY-1997: US-047581.  
 PR 23-MAY-1997: US-047582.  
 PR 23-MAY-1997: US-047583.  
 PR 23-MAY-1997: US-047584.  
 PR 23-MAY-1997: US-047585.  
 PR 23-MAY-1997: US-047586.  
 PR 23-MAY-1997: US-047587.  
 PR 23-MAY-1997: US-047588.  
 PR 23-MAY-1997: US-047589.  
 PR 23-MAY-1997: US-047590.  
 PR 23-MAY-1997: US-047592.  
 PR 23-MAY-1997: US-047593.  
 PR 23-MAY-1997: US-047594.  
 PR 23-MAY-1997: US-047595.  
 PR 23-MAY-1997: US-047596.  
 PR 23-MAY-1997: US-047597.  
 PR 23-MAY-1997: US-047598.  
 PR 23-MAY-1997: US-047599.  
 PR 23-MAY-1997: US-047600.  
 PR 23-MAY-1997: US-047601.  
 PR 23-MAY-1997: US-047612.  
 PR 23-MAY-1997: US-047613.  
 PR 23-MAY-1997: US-047614.  
 PR 23-MAY-1997: US-047615.  
 PR 23-MAY-1997: US-047617.  
 PR 23-MAY-1997: US-047618.  
 PR 23-MAY-1997: US-047632.  
 PR 23-MAY-1997: US-047633.  
 PR 06-JUN-1997: US-048964.  
 PR 06-JUN-1997: US-048974.  
 PR 13-JUN-1997: US-049610.  
 PR 08-JUL-1997: US-051926.

PR 16-JUL-1997: US-052874.  
 PR 18-AUG-1997: US-055724.  
 PR 22-AUG-1997: US-056630.  
 PR 22-AUG-1997: US-056631.  
 PR 22-AUG-1997: US-056632.  
 PR 22-AUG-1997: US-056636.  
 PR 22-AUG-1997: US-056637.  
 PR 22-AUG-1997: US-056662.  
 PR 22-AUG-1997: US-056664.  
 PR 22-AUG-1997: US-056845.  
 PR 22-AUG-1997: US-056862.  
 PR 22-AUG-1997: US-056864.  
 PR 22-AUG-1997: US-056872.  
 PR 22-AUG-1997: US-056874.  
 PR 22-AUG-1997: US-056875.  
 PR 22-AUG-1997: US-056876.  
 PR 22-AUG-1997: US-056877.  
 PR 22-AUG-1997: US-056878.  
 PR 22-AUG-1997: US-056879.  
 PR 22-AUG-1997: US-056880.  
 PR 22-AUG-1997: US-056881.  
 PR 22-AUG-1997: US-056882.  
 PR 22-AUG-1997: US-056884.  
 PR 22-AUG-1997: US-056886.  
 PR 22-AUG-1997: US-056887.  
 PR 22-AUG-1997: US-056888.  
 PR 22-AUG-1997: US-056889.  
 PR 22-AUG-1997: US-056892.  
 PR 22-AUG-1997: US-056893.  
 PR 22-AUG-1997: US-056894.  
 PR 22-AUG-1997: US-056903.  
 PR 22-AUG-1997: US-056908.  
 PR 22-AUG-1997: US-056909.  
 PR 22-AUG-1997: US-056910.  
 PR 22-AUG-1997: US-056911.  
 PR 05-SEP-1997: US-057650.  
 PR 05-SEP-1997: US-057669.  
 PR 05-SEP-1997: US-057761.  
 PR 12-SEP-1997: US-058785.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,  
 PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS,  
 PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,  
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;  
 DR WPI: 98-506364/43.  
 DR N-PSDB: V59743.  
 DR New isolated human genes and the secreted polypeptide(s) they encode  
 PT - useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 PS Claim 1; Page 674; 721pp; English.  
 PS This sequence represents a secreted human protein encoded by the nucleic  
 CC acid molecule designated Gene 82 from the human cDNA clone HPRAX55  
 CC (deposited as clone ATCC 97900 and ATCC 209046).  
 CC The gene can be used to generate fusion proteins by linking to the gene  
 CC to a human immunoglobulin Fc portion (e.g. V59502) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 186 novel genes and their fragments (nucleic  
 CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which  
 CC are useful for preventing, treating or ameliorating medical conditions  
 CC e.g. by protein or gene therapy. Also, pathological conditions can be  
 CC diagnosed by determining the amount of the new polypeptides in a sample  
 CC or by determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the 186 polynucleotides, based on  
 CC which tissues they are most highly expressed in (see V59511 for described  
 CC uses).  
 SQ Sequence 58 AA;  
 Query Match 14.3%; Score 5; DB 37; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 27 frsts 31  
 |||||

QY 28 FSRTS 32

## RESULT 12

ID W8625 standard; Protein; 59 AA.

AC W8625;

DT 01-MAR-1999 (first entry)

DE Secreted protein encoded by gene 92 clone HPCRI13.

KW Human; secreted protein; fusion protein; gene therapy; protein therapy;

KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;

KW developmental abnormality; foetal deficiency; blood; allergy; renal;

KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;

KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;

KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;

KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;

KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN W09854963-A2.

PD 10-DEC-1998.

PF 04-JUN-1998; U11422.

PR 18-DEC-1997; US-070923.

PR 06-JUN-1997; US-048877.

PR 06-JUN-1997; US-048881.

PR 06-JUN-1997; US-048893.

PR 06-JUN-1997; US-048896.

PR 06-JUN-1997; US-048899.

PR 06-JUN-1997; US-048915.

PR 06-JUN-1997; US-048949.

PR 06-JUN-1997; US-048964.

PR 06-JUN-1997; US-048972.

PR 06-JUN-1997; US-049020.

PR 06-JUN-1997; US-049375.

PR 05-SEP-1997; US-057628.

PR 05-SEP-1997; US-057635.

PR 05-SEP-1997; US-057644.

PR 05-SEP-1997; US-057647.

PR 05-SEP-1997; US-057650.

PR 05-SEP-1997; US-057661.

PR 05-SEP-1997; US-057667.

PR 05-SEP-1997; US-057761.

PR 05-SEP-1997; US-057764.

PR 05-SEP-1997; US-057770.

PR 05-SEP-1997; US-057775.

PR 05-SEP-1997; US-057778.

PR 06-JUN-1997; US-048875.

PR 06-JUN-1997; US-048878.

PR 06-JUN-1997; US-048882.

PR 06-JUN-1997; US-048885.

PR 06-JUN-1997; US-048894.

PR 06-JUN-1997; US-048897.

PR 06-JUN-1997; US-048900.

PR 06-JUN-1997; US-048916.

PR 06-JUN-1997; US-048962.

PR 06-JUN-1997; US-048970.

PR 06-JUN-1997; US-048974.

PR 06-JUN-1997; US-049373.

PR 05-SEP-1997; US-057584.

PR 05-SEP-1997; US-057629.

PR 05-SEP-1997; US-057642.

PR 05-SEP-1997; US-057645.

PR 05-SEP-1997; US-057648.

PR 05-SEP-1997; US-057651.

PR 05-SEP-1997; US-057662.

PR 05-SEP-1997; US-057668.

PR 05-SEP-1997; US-057762.

PR 05-SEP-1997; US-057765.

PR 05-SEP-1997; US-057771.

PR 05-SEP-1997; US-057776.

PR 06-JUN-1997; US-048876.

PR 06-JUN-1997; US-048880.

PR 06-JUN-1997; US-048883.

PR 06-JUN-1997; US-048892.

PR

06-JUN-1997; US-048895.

PR

06-JUN-1997; US-048898.

PR

06-JUN-1997; US-048901.

PR

06-JUN-1997; US-048917.

PR

06-JUN-1997; US-048963.

PR

06-JUN-1997; US-048971.

PR

06-JUN-1997; US-049019.

PR

06-JUN-1997; US-049374.

PR

05-SEP-1997; US-057627.

PR

05-SEP-1997; US-057634.

PR

05-SEP-1997; US-057643.

PR

05-SEP-1997; US-057646.

PR

05-SEP-1997; US-057649.

PR

05-SEP-1997; US-057654.

PR

05-SEP-1997; US-057666.

PR

05-SEP-1997; US-057760.

PR

05-SEP-1997; US-057763.

PR

05-SEP-1997; US-057769.

PR

05-SEP-1997; US-057774.

PR

05-SEP-1997; US-057777.

PA

(HUMA-) HUMAN GENOME SCI INC.

PI

Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA,

PI

Fan P, Feng P, Ferrie AM, Fischer CL, Florence C,

PI

Florence K, Greene JM, Hu J, Kyaw H, Laifleur DW,

PI

Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,

PI

Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;

DR

WPI: 99-059865/05.

DR

N-PSDB: V84502.

PT

New isolated human genes and the secreted polypeptides they encode -

PT

useful for diagnosis and treatment of e.g. cancers, neurological

PT

disorders, immune diseases, inflammation or blood disorders

PS

Claim 11: Page 530: 772pp: English.

CC

The invention relates to nucleic acid sequences (V84411 to V84633)

CC

encoding human secreted proteins (W88534 to W88756). The secreted protein

CC

gene sequences are deposited with the ATCC under deposit numbers ATCC

CC

97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,

CC

209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host

CC

cells comprising recombinant vectors containing the nucleic acid

CC

sequences are used for the recombinant production of the secreted

CC

proteins. The polynucleotide and amino acid sequences are useful for are

CC

useful for preventing, treating or ameliorating medical conditions e.g.

CC

by protein or gene therapy. Pathological conditions can be also diagnosed

CC

by determining the amount of the new polypeptides in a sample or by

CC

determining the presence of mutations in the new polynucleotides.

CC

Specific uses are described for each of the polynucleotides, based on

CC

which tissues they are most highly expressed in, and include developing

CC

products for the diagnosis or treatment of cancer, neurodegenerative

CC

disorders, developmental abnormalities and foetal deficiencies, blood

CC

disorders, tumours, leukemias, diseases of the immune system, autoimmune

CC

diseases, hepatic and renal disease, lymphomas, inflammation, allergies,

CC

ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,

CC

restenosis, prostate diseases, obesity, disorders involving osteoclasts

CC

such as osteoporosis, arthritis or malignancies, diseases of testes, lung

CC

or thymus, digestive/endocrine disorders, infections and AIDS. The

CC

polypeptides are also useful for identifying their binding partners.

CC

The present sequence represents human secreted protein (see descriptor

CC

line for gene number and clone identification).

SQ

Sequence 59 AA;

Query Match 14.3%; Score 5; DB 38; Length 59;

Best Local Similarity 100.0%; Pred. No. 3.25e-02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 frsts 6

QY 28 FSRTS 32

## RESULT 13

ID W55449 standard; Protein; 91 AA.

AC W55449;

DT 24-JUN-1998 (first entry)

DE H. pylori ORF 02ae11612\_1074212\_f1\_1 secreted protein.



KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
 KW identification; binding compound; bacteria; life cycle; activator;  
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.  
 OS Helicobacter pylori.  
 PN W09737044-AL.  
 PD 09-OCT-1997.  
 PF 27-MAR-1997; U05223.  
 PR 06-DEC-1996; US-761318.  
 PR 29-MAR-1996; US-625811.  
 PR 02-APR-1996; US-758731.  
 PR 25-OCT-1996; US-736905.  
 PR 28-OCT-1996; US-738859.  
 PA (ASTR.) ASTRA AB.  
 PI Alm RA, Smith D;  
 DR WPI: 97-503122/46.  
 DR N-PSDB; V24858.  
 PT Helicobacter pylori nucleic acid sequences and encoded  
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori  
 PT infection and for diagnosis of H. pylori infection  
 PS Claims 14,94; Page 656; 1145pp; English.  
 CC This sequence is a H. pylori secreted protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors. The  
 CC DNA and probes derived from it may be used for the identification of  
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic  
 CC acid sequences complementary to the DNA act as antisense sequences and  
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies  
 CC against the protein can be used in immunoassays to evaluate the abundance  
 CC and distribution of H. pylori-specific antigens. The genomic sequence of  
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated  
 CC by mechanically shearing the bacterial DNA. The sequences were analysed  
 CC for ORF of at least 180 nucleotides, and the predicted coding regions  
 CC defined by computer evaluation. To identify likely H. pylori antigens for  
 CC vaccine development, the amino acid sequences predicted from various ORF  
 CC were analysed for significant homology to other known or exported  
 CC membrane proteins. Having identified and determined the sequences of  
 CC interest, particular regions can be isolated from H. pylori by PCR  
 CC amplification for recombinant polypeptide production, e.g. in E. coli  
 CC hosts.  
 SQ Sequence 91 AA;

Query Match 14.3%; Score 5; DB 29; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 endei 11  
 QY 3 ENDEI 7

RESULT 14  
 ID W27984 standard; Protein; 95 AA.  
 AC W27984;  
 DT 26-AUG-1998 (first entry)  
 DE Staphylococcus aureus protein of unknown function.  
 KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;  
 KW Staphylococcal gene; regulatory element; bacterial gene expression;  
 KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;  
 KW toxic shock syndrome.  
 OS Staphylococcus aureus.  
 PN W09730070-AL.  
 PD 21-AUG-1997.  
 PF 19-FEB-1997; U02318.  
 PR 20-FEB-1996; US-011888.  
 PA (SMIK.) SMITHKLINE BEECHAM CORP.  
 PI Black Mt, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,  
 PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;  
 DR WPI: 97-424969/39.  
 DR N-PSDB; T83943.  
 PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used  
 PT to isolate antimicrobial compounds, and in vaccines against S.  
 PT aureus infection

PS Claim 6: Page 389; 989pp; English.  
 CC The present sequence represents a Staphylococcus aureus protein of  
 CC unknown function. The DNA sequence was isolated from a library of  
 CC clones of S. aureus WCUH 29 in Escherichia coli. The DNA sequence can  
 CC be used in the construction of ribozymes and antisense sequences to  
 CC control the expression of staphylococcal genes. The DNA sequence is  
 CC also useful as a source of regulatory elements for the control of  
 CC bacterial gene expression. The present protein may be used to produce  
 CC vaccines to enable a host to produce specific antibodies with  
 CC antibacterial action. These vaccines and antibodies would protect  
 CC a host against invasion by S. aureus, and conditions relating to  
 CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled  
 CC skin syndrome, and toxic shock syndrome.  
 SQ Sequence 95 AA;

Query Match 14.3%; Score 5; DB 31; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 82 iende 86  
 QY 2 IENDE 6

RESULT 15  
 ID P81777 standard; protein; 99 AA.  
 AC P81777;  
 DT 15-NOV-1990 (first entry)  
 DE Sequence encoded by open reading frame of cDNA corresponding to  
 DE HIV-2 ROD genome  
 DE LAV-II ROD; AIDS; immunogen; antigen; vaccine; diagnostic.  
 KW LAV-II ROD; AIDS; immunogen; antigen; vaccine; diagnostic.  
 OS Human immunodeficiency virus ROD.  
 PN W08805440-A.  
 PD 28-JUL-1988.  
 PR 15-JAN-1988;  
 PR 15-APR-1987; FR-005398.  
 PA (INSP) Inst Pasteur(ALIZ/).  
 PI Allison M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M,  
 PI Tiollais P, Chakrabarti L, Desrosiers R;  
 DR WPI: 88-220290/31.  
 DR N-PSDB; n80859.  
 PT New peptide(s) with immunological properties of HIV-2 envelope protein -  
 PT having the structure of simian immune deficiency virus proteins,  
 PT useful in diagnosis and of vaccine components  
 PS Disclosure; Fig 1A; 86pp; French.  
 CC The SQ in n80859 was deposited on 21/2/86 at the CNMCM under number I-522.  
 CC reference name LAV-II ROD. It is the cDNA to HIV-2 ROD genomic RNA. The SQ  
 CC was compared with the SQ of the genome of SIV (Mac) (n80860) to identify  
 CC common regions.  
 SQ Sequence 99 AA;

Query Match 14.3%; Score 5; DB 1; Length 99;  
 Best Local Similarity 100.0%; Pred. No. 3.25e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 19 fsrts 23  
 QY 28 FSRTS 32

Search completed: Sat Aug 28 15:09:56 1999  
 Job time : 76 secs.

This Page Blank (usps)

\*\*\*\*\*  
WQSEH  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:07:34 1999; MasPar time 5.18 Seconds  
Tabular output not generated.

Title: >US-09-049-696-48  
Description: (1-35) from US09049696.ppe  
Sequence: 1 WIENDEIQWNPPEINKDDVQKQVCFSTSGG 35

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 122810 seqs, 40068593 residues  
Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: pir60  
1.pir1 2.pir2 3.pir3 4.pir4

Statistics: Mean 2.688; Variance 0.419; scale 6.423

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result	Query	Description	Pred. No.
No.	Score	Match Length DB ID	
1	7	20.0	287 2 S24800
2	6	17.1	131 2 T03279
3	6	17.1	183 2 A24479
4	6	17.1	183 2 A24479
5	6	17.1	184 2 A1230
6	6	17.1	185 2 S13183
7	6	17.1	185 2 S13184
8	6	17.1	185 2 S13185
9	6	17.1	185 2 S13182
10	6	17.1	334 2 E70079
11	6	17.1	370 2 A25004
12	6	17.1	423 2 F69031
13	6	17.1	440 2 B1153
14	6	17.1	483 2 A34720
15	6	17.1	505 1 E69417
16	6	17.1	574 2 E64414
17	6	17.1	671 2 A40692
18	6	17.1	794 2 T02171
19	6	17.1	820 2 T02152
20	6	17.1	905 2 T02205
21	6	17.1	913 3 JG0168
22	6	17.1	916 2 S2572
23	6	17.1	1680 2 T01367

Clastrin heavy chain 5.36e+00  
hypothetical protein, 5.36e+00  
genome polyprotein - 5.36e+00  
nonstructural protein - 5.36e+00  
nonstructural polypro 5.36e+00  
notch protein - fruit 5.36e+00  
ankyrin 3, long splic 5.36e+00  
calcium-dependent cha 1.95e+02  
hypothetical protein 1.95e+02  
hypothetical protein, 1.95e+02  
hypothetical protein, 1.95e+02  
hypothetical protein 1.95e+02  
hypothetical protein 1.95e+02  
hypothetical 7K prote 1.95e+02  
homeotic protein 2.7 1.95e+02  
hypothetical 9K prote 1.95e+02  
trans-regulatory spli 1.95e+02  
hypothetical protein 1.95e+02  
neuropeptide y precur 1.95e+02  
Ig heavy chain V-VI r 1.95e+02  
E7 protein - human pa 1.95e+02  
ferredoxin precursor 1.95e+02  
conserved hypotheticala 1.95e+02  
hypothetical protein 1.95e+02  
adenine phosphoriboxy 1.95e+02  
Ig heavy chain V regi 1.95e+02  
hypothetical protein 1.95e+02  
A-ORF-A protein - vac 1.95e+02  
bombersinlike peptide 1.95e+02  
hypothetical protein 1.95e+02  
trans-activating tran 1.95e+02  
probable membrane pro 1.95e+02  
hypothetical protein 1.95e+02  
yhcV homolog - Methan 1.95e+02  
T-cell receptor alpha 1.95e+02  
H-transporing ATP s 1.95e+02  
hypothetical protein 1.95e+02  
conserved hypotheticala 1.95e+02  
coat protein 2 - bean 1.95e+02  
ribosomal protein L27 1.95e+02  
ribosomal protein L27 1.95e+02  
crk protein - Rhodob 1.95e+02  
tryptophan-rich senso 1.95e+02  
MADS box protein - ma 1.95e+02  
T-cell receptor delta 1.95e+02  
ZEM2 protein - maize 1.95e+02  
probable NADH-ubiquin 1.95e+02  
hypothetical protein 1.95e+02  
hypothetical protein 1.95e+02  
probable membrane pro 1.95e+02  
precorrin-8W decarbox 1.95e+02  
probable nuob protein 1.95e+02  
conserved hypotheticala 1.95e+02  
conserved hypotheticala 1.95e+02  
hypothetical protein 1.95e+02  
chitin synthase (EC 2 1.95e+02  
adenylate kinase (EC 1.95e+02  
proline-rich protein 1.95e+02  
conserved hypotheticala 1.95e+02  
hypothetical protein 1.95e+02  
hypothetical protein 1.95e+02  
incB protein - phage 1.95e+02  
ZEM1 protein - maize 1.95e+02  
orf3 - Southampton vi 1.95e+02  
coat protein - pea ea 1.95e+02  
grp cyclohydrolase I 1.95e+02  
hypothetical protein 1.95e+02  
probable lppl protein 1.95e+02  
ribosomal protein S3, 1.95e+02  
hypothetical ndhk hom 1.95e+02  
triose-phosphate isom 1.95e+02

97	5	14.3	220	1	E65000	NADH dehydrogenase (u	1.95e+02	170	327	2	A46484	apoptosis-mediating m	1.95e+02
98	5	14.3	220	2	J01985	avrPial protein - Ps	1.95e+02	171	327	2	T00797	hypothetical protein	1.95e+02
99	5	14.3	227	2	C39925	hypothetical protein	1.95e+02	172	328	2	B64095	DNA-directed RNA poly	1.95e+02
100	5	14.3	228	2	C26599	clathrin light chain	1.95e+02	173	329	2	S58158	hypothetical protein	1.95e+02
101	5	14.3	228	2	S25730	sem-5 protein - Caeno	1.95e+02	174	330	3	JE0376	Grb-2 related adaptor	1.95e+02
102	5	14.3	229	2	C70164	rRNA methylase (yaco)	1.95e+02	175	331	2	E65076	hypothetical protein	1.95e+02
103	5	14.3	229	2	B31775	clathrin light chain	1.95e+02	176	332	2	A29711	deactoxycyphalospori	1.95e+02
104	5	14.3	230	1	WNV26	p26 protein - Orgvla	1.95e+02	177	333	2	B49850	protochlorophyllide r	1.95e+02
105	5	14.3	231	2	B69477	ABC transporter, ATP-	1.95e+02	178	333	2	S17823	protochlorophyllide r	1.95e+02
106	5	14.3	231	2	S28703	hypothetical protein	1.95e+02	179	333	2	S30915	protochlorophyllide r	1.95e+02
107	5	14.3	232	2	C70699	probable pABA protein	1.95e+02	180	339	2	A34365	lic-1 phase variation	1.95e+02
108	5	14.3	238	2	I52638	neuronal growth-relat	1.95e+02	181	339	2	A42259	endo-beta-N-acetylglu	1.95e+02
109	5	14.3	238	2	D71706	hypothetical protein	1.95e+02	182	340	2	H0790	hypothetical protein	1.95e+02
110	5	14.3	239	1	WZBED5	26.5K membrane protei	1.95e+02	183	343	2	H64491	hypothetical protein	1.95e+02
111	5	14.3	239	2	T01406	bovine corneal protei	1.95e+02	184	343	1	CES0VM	vancomycin resistance	1.95e+02
112	5	14.3	239	2	S64327	probable membrane pro	1.95e+02	185	345	1	NEMS	apolipoprotein H prec	1.95e+02
113	5	14.3	240	2	H70045	hypothetical protein	1.95e+02	186	346	2	S44874	2C21.3 protein - Caen	1.95e+02
114	5	14.3	242	2	T02193	hypothetical protein	1.95e+02	187	346	2	C65066	hypothetical protein	1.95e+02
115	5	14.3	243	2	E64452	conserved hypotheticala	1.95e+02	188	347	1	QOCVS1	Al1 protein - squash	1.95e+02
116	5	14.3	243	2	A26599	clathrin light chain	1.95e+02	189	349	2	H61162	conserved hypotheticala	1.95e+02
117	5	14.3	243	2	S74216	NADH dehydrogenase (u	1.95e+02	190	354	2	G71403	hypothetical protein	1.95e+02
118	5	14.3	248	1	S04437	NADH dehydrogenase (u	1.95e+02	191	355	2	I64126	hypothetical protein	1.95e+02
119	5	14.3	248	2	A31775	clathrin light chain	1.95e+02	192	355	2	S56058	probable membrane pro	1.95e+02
120	5	14.3	248	2	T02647	probable DNA-binding	1.95e+02	193	360	2	E70816	probable moaA2 protei	1.95e+02
121	5	14.3	250	2	PC1117	GTP cyclohydrolase I	1.95e+02	194	361	2	S30325	hypothetical protein	1.95e+02
122	5	14.3	251	2	T03298	expansin 2 - rice	1.95e+02	195	362	2	C64807	ybG0 protein - Escher	1.95e+02
123	5	14.3	252	2	E69034	hypothetical protein	1.95e+02	196	365	2	T00599	hypothetical protein	1.95e+02
124	5	14.3	255	2	S70545	1'-acetylglucosyl-3-phos	1.95e+02	197	372	2	S39559	mitogen-activated pro	1.95e+02
125	5	14.3	255	2	T02010	expansin homolog T15B	1.95e+02	198	373	2	S17955	luxE protein - Photob	1.95e+02
126	5	14.3	257	2	T02727	probable expansin T91	1.95e+02	199	373	2	JH0155	pectin lyase (EC 4.2.	1.95e+02
127	5	14.3	257	2	S70544	1'-acetylglucosyl-3-phos	1.95e+02	200	380	2	JQ1679	DNA-binding protein G	1.95e+02
128	5	14.3	258	2	E71421	hypothetical protein	1.95e+02	201	382	2	S51962	FUN4 protein - yeast	1.95e+02
129	5	14.3	258	2	S53082	pollen allergen homol	1.95e+02	202	383	2	S36552	E2 protein - human pa	1.95e+02
130	5	14.3	263	2	E64860	hypothetical protein	1.95e+02	203	384	2	S52989	mitogen-activated, ex	1.95e+02
131	5	14.3	263	2	S73314	hypothetical protein	1.95e+02	204	384	2	H64161	hypothetical protein	1.95e+02
132	5	14.3	263	2	C70583	probable dehydrogenas	1.95e+02	205	385	2	I84455	dopamine transporter	1.95e+02
133	5	14.3	265	2	A56557	homeotic protein eng-	1.95e+02	206	385	2	JC5019	aldehyde dehydrogenas	1.95e+02
134	5	14.3	267	2	E64172	hypothetical protein	1.95e+02	207	386	2	S14874	ADP-ATP carrier prote	1.95e+02
135	5	14.3	272	1	S28795	carbonate dehydratase	1.95e+02	208	387	2	S64082	probable membrane pro	1.95e+02
136	5	14.3	275	2	A32057	nitrogenase (EC 1.18.	1.95e+02	209	389	2	S39781	phosphoprotein phosph	1.95e+02
137	5	14.3	276	2	E71705	hypothetical protein	1.95e+02	210	389	2	S44643	F37A4.6 protein - Cae	1.95e+02
138	5	14.3	279	2	S73526	probable lipoprotein	1.95e+02	211	390	2	S39780	phosphoprotein phosph	1.95e+02
139	5	14.3	281	2	E64306	hypothetical protein	1.95e+02	212	390	2	S20392	phosphoprotein phosph	1.95e+02
140	5	14.3	281	1	PNSAP	beta-lactamase (EC 3.	1.95e+02	213	390	2	S65672	phosphoprotein phosph	1.95e+02
141	5	14.3	284	2	H69355	conserved hypotheticala	1.95e+02	214	390	2	B70011	NADH-dependent butano	1.95e+02
142	5	14.3	285	2	C70873	probable enoyl-CoA hy	1.95e+02	215	391	2	S74720	DNA-directed DNA poly	1.95e+02
143	5	14.3	286	1	LRR7A1	clathrin light chain	1.95e+02	216	392	2	PQ0446	DNA-binding protein B	1.95e+02
144	5	14.3	289	2	S70293	hypothetical protein	1.95e+02	217	392	2	S37902	hypothetical protein	1.95e+02
145	5	14.3	294	2	S65160	APG5 protein - yeast	1.95e+02	218	392	2	S22580	telomere-binding prot	1.95e+02
146	5	14.3	297	2	S64308	hypothetical protein	1.95e+02	219	393	2	I49016	phosphoprotein phosph	1.95e+02
147	5	14.3	300	2	E64308	acetylglutamate kinas	1.95e+02	220	396	2	A71206	hypothetical protein	1.95e+02
148	5	14.3	303	2	G71166	hypothetical protein	1.95e+02	221	397	2	JC2524	phosphoprotein phosph	1.95e+02
149	5	14.3	304	2	S61033	hypothetical protein	1.95e+02	222	402	1	VCBE50	glycoprotein D precur	1.95e+02
150	5	14.3	307	2	S56317	rpI protein - Escher	1.95e+02	223	402	2	I67760	transposase - Escher	1.95e+02
151	5	14.3	308	2	I36791	hypothetical protein	1.95e+02	224	403	2	S28144	acetyl-CoA C-acetyltr	1.95e+02
152	5	14.3	308	2	E71697	probable proteinase s	1.95e+02	225	403	2	T02003	probable DNA-binding	1.95e+02
153	5	14.3	309	2	S78391	probable transcriptio	1.95e+02	226	406	2	A47492	phosphoprotein phosph	1.95e+02
154	5	14.3	311	2	A32043	deactoxycyphalospori	1.95e+02	227	408	2	S73681	nitrogen fixation pro	1.95e+02
155	5	14.3	314	2	S30900	deactoxycyphalospori	1.95e+02	228	409	2	S44346	RAD23 protein homolog	1.95e+02
156	5	14.3	314	2	C70562	probable DTRP-Glucose	1.95e+02	229	409	2	B25156	cellulase (EC 3.2.1.4	1.95e+02
157	5	14.3	314	2	T01554	hypothetical protein	1.95e+02	230	409	2	S18984	arrestin - human (fra	1.95e+02
158	5	14.3	315	2	S32401	H+-transporting Atp s	1.95e+02	231	410	2	A43404	beta-arrestin2 - rat	1.95e+02
159	5	14.3	316	2	S77783	hypothetical protein	1.95e+02	232	412	2	S50963	hypothetical protein	1.95e+02
160	5	14.3	317	2	G64408	deactoxycyphalosporin	1.95e+02	233	413	2	T02693	S-receptor kinase hom	1.95e+02
161	5	14.3	319	2	S54100	hypothetical protein	1.95e+02	234	416	1	A41267	transcription factor	1.95e+02
162	5	14.3	321	2	S59416	hypothetical protein	1.95e+02	235	416	2	A32947	filaggrin precursor -	1.95e+02
163	5	14.3	321	2	B69632	glucose kinase glck -	1.95e+02	236	417	1	TVCRCG	phosphoglycerate kina	1.95e+02
164	5	14.3	322	2	E70326	conserved hypotheticala	1.95e+02	237	418	2	A64508	hypothetical protein	1.95e+02
165	5	14.3	322	2	E71137	hypothetical protein	1.95e+02	238	419	2	S23425	preprotein translocas	1.95e+02
166	5	14.3	324	2	A36959	host-specific nodulat	1.95e+02	239	420	2	A47140	arrestin arr3L - bovi	1.95e+02
167	5	14.3	325	2	G70334	alkalinal monooxygenase	1.95e+02	240	420	1	TVUFG4	phosphoglycerate kina	1.95e+02
168	5	14.3	326	2	S17954	hypothetical protein	1.95e+02	241	421	1	KIUTGC	phosphoglycerate kina	1.95e+02
169	5	14.3	327	2	S22298	CAMP response element	1.95e+02	242	421	2	S70402	zona pellucida glycop	1.95e+02

243	5	14.3	422	2	A34589	lin-10 protein - Caen	1.95e+02	316	5	14.3	536	2	D42453	hypothetical protein	1.95e+02
244	5	14.3	424	1	S07792	site-specific DNA-met	1.95e+02	317	5	14.3	538	2	G09317	conserved hypothetical	1.95e+02
245	5	14.3	427	1	G8BP24	gene 24 protein - pha	1.95e+02	318	5	14.3	539	2	A54294	cell division control	1.95e+02
246	5	14.3	427	2	D64813	ybcH protein precursor	1.95e+02	319	5	14.3	539	2	S67049	probable membrane pro	1.95e+02
247	5	14.3	430	2	A56002	pre-B-cell leukemia t	1.95e+02	320	5	14.3	539	2	A42467	alkaline phosphatase	1.95e+02
248	5	14.3	430	2	JC2301	hypothetical 47'8K pr	1.95e+02	321	5	14.3	544	2	C42653	dihydrolipoamide S-ac	1.95e+02
249	5	14.3	432	2	S41956	rad9 protein - Schizo	1.95e+02	322	5	14.3	547	2	A43832	heme-binding protein	1.95e+02
250	5	14.3	434	2	S51644	secreted/adhesive pro	1.95e+02	323	5	14.3	549	2	D64098	heme-binding protein	1.95e+02
251	5	14.3	435	2	A44308	Antho-RfamId precurs	1.95e+02	324	5	14.3	549	2	G71606	hypothetical protein	1.95e+02
252	5	14.3	440	1	TVUTGB	phosphoglycerate kina	1.95e+02	325	5	14.3	552	2	C70567	probable ilvB2 protei	1.95e+02
253	5	14.3	440	1	KUTGG	phosphoglycerate kina	1.95e+02	326	5	14.3	552	2	S55026	secretion protein - A	1.95e+02
254	5	14.3	440	2	D64616	N-acetylmuramoyl-L-al	1.95e+02	327	5	14.3	556	2	S51892	probable membrane pro	1.95e+02
255	5	14.3	448	1	A60003	nucleocapsid protein	1.95e+02	328	5	14.3	557	2	A25493	tryptophan 2-monooxyg	1.95e+02
256	5	14.3	449	2	I40178	probable glutathione	1.95e+02	329	5	14.3	565	2	F70152	hypothetical protein	1.95e+02
257	5	14.3	451	2	T00694	hypothetical protein	1.95e+02	330	5	14.3	567	2	A45977	Rab geranylgeranyl tr	1.95e+02
258	5	14.3	452	2	T00840	physical impedance pr	1.95e+02	331	5	14.3	569	2	C69471	probable fatty-acid--	1.95e+02
259	5	14.3	453	2	S56657	aspartate transaminas	1.95e+02	332	5	14.3	573	2	E70169	phosphoenolpyruvate-p	1.95e+02
260	5	14.3	453	2	A30149	aldehyde dehydrogenas	1.95e+02	333	5	14.3	579	2	B45266	MPL-K protein precurs	1.95e+02
261	5	14.3	453	2	A42584	aldehyde dehydrogenas	1.95e+02	334	5	14.3	581	2	C71078	probable aldehyde:fer	1.95e+02
262	5	14.3	453	2	S40943	hypothetical protein	1.95e+02	335	5	14.3	584	2	S55106	probable membrane pro	1.95e+02
263	5	14.3	454	2	A64305	conserved hypothetical	1.95e+02	336	5	14.3	590	1	TVFFDS	protein-tyrosine kina	1.95e+02
264	5	14.3	454	1	C69214	zinc metalloprotein	1.95e+02	337	5	14.3	595	1	S20825	protein-tyrosine-phos	1.95e+02
265	5	14.3	455	1	TVCRGG	phosphoglycerate kina	1.95e+02	338	5	14.3	595	1	A44390	protein-tyrosine-phos	1.95e+02
266	5	14.3	456	2	S19212	adenylosuccinate lyas	1.95e+02	339	5	14.3	597	2	S72458	probable transcriptio	1.95e+02
267	5	14.3	460	2	JC2194	vasoactive intestinal	1.95e+02	340	5	14.3	599	2	T00659	glucose-6-phosphate 1	1.95e+02
268	5	14.3	460	2	B60048	cyclin B4 - yeast (Sa	1.95e+02	341	5	14.3	607	2	T01904	hypothetical protein	1.95e+02
269	5	14.3	461	2	S36593	E2 protein - human pa	1.95e+02	342	5	14.3	609	2	S67616	probable membrane pro	1.95e+02
270	5	14.3	463	2	S37962	probable purine nucle	1.95e+02	343	5	14.3	609	2	S28283	hypothetical protein	1.95e+02
271	5	14.3	466	2	A39599	55k erythrocyte membr	1.95e+02	344	5	14.3	611	2	S38162	translation elongatio	1.95e+02
272	5	14.3	468	1	FQ808	fibrinogen beta chain	1.95e+02	345	5	14.3	615	2	T03027	receptor-like protein	1.95e+02
273	5	14.3	468	2	I38669	ALDH7 - human	1.95e+02	346	5	14.3	617	2	T02668	probable receptor ser	1.95e+02
274	5	14.3	468	2	D59716	involved in spore cor	1.95e+02	347	5	14.3	618	2	F70855	probable ilvB protein	1.95e+02
275	5	14.3	469	2	H71897	probable probable n-a	1.95e+02	348	5	14.3	619	2	I59558	dopamine transporter	1.95e+02
276	5	14.3	471	2	S05392	hypothetical protein	1.95e+02	349	5	14.3	619	2	S20346	dopamine transporter pr	1.95e+02
277	5	14.3	475	2	S46663	ribulose-bisphosphate	1.95e+02	350	5	14.3	620	2	A48980	dopamine transporter	1.95e+02
278	5	14.3	476	2	A44170	membrane-bound ribos	1.95e+02	351	5	14.3	620	2	I57937	dopamine transporter	1.95e+02
279	5	14.3	477	2	I51645	stromelysin 3 (EC 3.4	1.95e+02	352	5	14.3	621	2	JC5164	acetolactate synthase	1.95e+02
280	5	14.3	478	2	S1457	probable membrane pro	1.95e+02	353	5	14.3	622	2	T02244	probable DNA repair p	1.95e+02
281	5	14.3	479	1	VGBEPB	glycoprotein gIII pre	1.95e+02	354	5	14.3	623	2	A49840	segment polarity prot	1.95e+02
282	5	14.3	483	2	S40128	fascin - African claw	1.95e+02	355	5	14.3	630	2	T02524	probable zinc finger	1.95e+02
283	5	14.3	483	2	D31266	aldehyde dehydrogenas	1.95e+02	356	5	14.3	631	2	E69115	phage infection prote	1.95e+02
284	5	14.3	484	2	A41028	aldehyde dehydrogenas	1.95e+02	357	5	14.3	635	2	A45266	MPL-P protein precurs	1.95e+02
285	5	14.3	487	2	S44265	glycoprotein gc - bov	1.95e+02	358	5	14.3	635	2	S76371	hypothetical protein	1.95e+02
286	5	14.3	488	2	A35156	cellulase (EC 3.2.1.4	1.95e+02	359	5	14.3	636	2	C90025	dnak-type molecular c	1.95e+02
287	5	14.3	490	2	S52830	HMS1 protein - yeast	1.95e+02	360	5	14.3	637	2	E70887	probable fadB32 prote	1.95e+02
288	5	14.3	493	2	S46283	calcium-dependent pro	1.95e+02	361	5	14.3	639	2	A39135	DNA topoisomerase cha	1.95e+02
289	5	14.3	495	2	JC2195	vasoactive intestinal	1.95e+02	362	5	14.3	641	2	D71170	hypothetical protein	1.95e+02
290	5	14.3	496	2	B45868	glycerol kinase (EC 2	1.95e+02	363	5	14.3	643	2	S60710	alpha 1,2 mannosidase	1.95e+02
291	5	14.3	496	2	S33791	ARS-binding protein -	1.95e+02	364	5	14.3	647	2	F70057	penicillin-binding pr	1.95e+02
292	5	14.3	497	2	A53373	rpoN protein - pseudo	1.95e+02	365	5	14.3	654	2	S76870	hypothetical protein	1.95e+02
293	5	14.3	500	2	A43656	sigK-creating site-sp	1.95e+02	366	5	14.3	657	2	S51387	dnak-type molecular c	1.95e+02
294	5	14.3	501	2	S54576	probable membrane pro	1.95e+02	367	5	14.3	657	2	H71422	hypothetical protein	1.95e+02
295	5	14.3	503	2	J01019	site-specific DNA-met	1.95e+02	368	5	14.3	658	1	SVBYDM	aspartate--tRNA ligas	1.95e+02
296	5	14.3	503	1	CBPRH	site-specific DNA-met	1.95e+02	369	5	14.3	661	2	H64876	probable membrane pro	1.95e+02
297	5	14.3	505	2	F71218	hypothetical protein	1.95e+02	370	5	14.3	662	2	T01533	hypothetical protein	1.95e+02
298	5	14.3	505	1	TVUT2B	phosphoglycerate kina	1.95e+02	371	5	14.3	664	2	S60757	hypothetical protein	1.95e+02
299	5	14.3	505	1	KICRFG	phosphoglycerate kina	1.95e+02	372	5	14.3	667	2	S60709	alpha 1,2 mannosidase	1.95e+02
300	5	14.3	506	1	ACRYGL	nicotinic acetylcholi	1.95e+02	373	5	14.3	674	2	S78570	probable membrane pro	1.95e+02
301	5	14.3	507	2	D64083	adenylosuccinate lyas	1.95e+02	374	5	14.3	676	2	F69276	conserved hypothetical	1.95e+02
302	5	14.3	508	1	TVUT4B	phosphoglycerate kina	1.95e+02	375	5	14.3	676	2	S66706	translation initiatio	1.95e+02
303	5	14.3	508	2	S44264	glycoprotein gc - bov	1.95e+02	376	5	14.3	684	2	A56154	Abi substrate ena (en	1.95e+02
304	5	14.3	509	1	A45593	phosphoglycerate kina	1.95e+02	377	5	14.3	686	2	JC5708	villin-like protein -	1.95e+02
305	5	14.3	511	2	S55198	sulfate adenylyltrans	1.95e+02	378	5	14.3	686	1	ALDYAT	amylase A (EC 3.2.1.-	1.95e+02
306	5	14.3	511	1	VGVN	spike glycoprotein G	1.95e+02	379	5	14.3	691	2	S48390	hypothetical protein	1.95e+02
307	5	14.3	513	2	S63701	mannosyl-oligosacchar	1.95e+02	380	5	14.3	691	1	VGNVH3	capsid-associated pro	1.95e+02
308	5	14.3	521	2	H07039	probable polyketide s	1.95e+02	381	5	14.3	692	2	S47734	hypothetical protein	1.95e+02
309	5	14.3	521	1	VGBEBB	glycoprotein gIII pre	1.95e+02	382	5	14.3	698	2	D64084	helicase (EC 3.6.1.-)	1.95e+02
310	5	14.3	526	2	S52190	clostripain (EC 3.4.2	1.95e+02	383	5	14.3	704	2	S39004	finger protein MSN2 -	1.95e+02
311	5	14.3	531	2	S52294	2-isopropylmalate syn	1.95e+02	384	5	14.3	704	2	A48040	meprin A (EC 3.4.24.1	1.95e+02
312	5	14.3	532	1	RSVOC2	RNA-directed RNA poly	1.95e+02	385	5	14.3	709	2	S75212	comeE protein - Synsch	1.95e+02
313	5	14.3	533	2	J50304	developmental control	1.95e+02	386	5	14.3	718	2	G71888	flagellar hook protei	1.95e+02
314	5	14.3	534	2	S61185	hypothetical protein	1.95e+02	387	5	14.3	718	2	F64628	flagellar hook protei	1.95e+02
315	5	14.3	536	2	I67945	phosphodiesterase - r	1.95e+02	388	5	14.3	718	1	ALBSMX	cyclomaltoextrin glu	1.95e+02

389	5	14.3	722	2	S64492	hypothenical protein	1.95e+02	462	5	14.3	1096	2	S61917	protein kinase C (EC	1.95e+02
390	5	14.3	727	2	A33552	NADH dehydrogenase (u	1.95e+02	463	5	14.3	1097	2	S47320	protein kinase C (EC	1.95e+02
391	5	14.3	732	2	S23001	tra1 protein - Escher	1.95e+02	464	5	14.3	1102	2	S42247	virg protein - Shigel	1.95e+02
392	5	14.3	732	2	S69749	hypothenical protein	1.95e+02	465	5	14.3	1114	2	I50222	deltaE1 - chicken	1.95e+02
393	5	14.3	732	2	S44617	C50C3.11 protein - Ca	1.95e+02	466	5	14.3	1118	2	S44641	F37A4.4 protein - cae	1.95e+02
394	5	14.3	739	2	F71161	probable chemotaxis p	1.95e+02	467	5	14.3	1120	2	T01863	hypothenical protein	1.95e+02
395	5	14.3	748	2	S59327	hypothenical protein	1.95e+02	468	5	14.3	1139	2	S61918	protein kinase C (EC	1.95e+02
396	5	14.3	758	2	S36711	dolichyl-phosphate-ma	1.95e+02	469	5	14.3	1146	2	S64402	probable transcriptio	1.95e+02
397	5	14.3	759	2	A35597	SEC2 protein - yeast	1.95e+02	470	5	14.3	1173	2	A53430	1-phosphatidylinosito	1.95e+02
398	5	14.3	781	2	S51592	XynB precursor - Rumi	1.95e+02	471	5	14.3	1174	2	A40853	potassium channel pro	1.95e+02
399	5	14.3	790	2	S27458	SMR2 protein - yeast	1.95e+02	472	5	14.3	1181	2	A43346	1-phosphatidylinosito	1.95e+02
400	5	14.3	795	2	JC4234	gene fused protein -	1.95e+02	473	5	14.3	1199	2	A40670	nuclear envelope prot	1.95e+02
401	5	14.3	810	2	S69652	hypothenical protein	1.95e+02	474	5	14.3	1208	2	T00362	hypothenical protein	1.95e+02
402	5	14.3	810	2	A44054	orf1 protein - Junoni	1.95e+02	475	5	14.3	1211	2	S68251	phospholipase C, inos	1.95e+02
403	5	14.3	827	2	A37849	S-layer protein - Hal	1.95e+02	476	5	14.3	1216	2	A28821	1-phosphatidylinosito	1.95e+02
404	5	14.3	828	2	D71700	leucyl-tRNA synthetas	1.95e+02	477	5	14.3	1216	2	A28822	1-phosphatidylinosito	1.95e+02
405	5	14.3	835	2	S11442	invasin - Versinia en	1.95e+02	478	5	14.3	1222	2	A28977	hypothenical protein	1.95e+02
406	5	14.3	835	2	S54216	invasin - Versinia en	1.95e+02	479	5	14.3	1224	2	S73171	DNA-directed RNA poly	1.95e+02
407	5	14.3	843	2	S53947	hypothenical protein	1.95e+02	480	5	14.3	1228	2	S47514	hypothenical protein	1.95e+02
408	5	14.3	844	2	T00529	bZIP-like protein - A	1.95e+02	481	5	14.3	1242	1	DJBEC1	DNA-directed DNA poly	1.95e+02
409	5	14.3	845	1	GLVK	beta-glucosidase (EC	1.95e+02	482	5	14.3	1252	2	D71810	probable type II DNA	1.95e+02
410	5	14.3	847	2	A53800	mixed-lineage protein	1.95e+02	483	5	14.3	1253	2	S46949	arylphorin receptor -	1.95e+02
411	5	14.3	852	2	A41886	DNA repair protein -	1.95e+02	484	5	14.3	1259	2	JC5523	transmembrane protein	1.95e+02
412	5	14.3	861	2	T00434	Kinesin heavy chain h	1.95e+02	485	5	14.3	1269	2	S35366	furl protein - fruit	1.95e+02
413	5	14.3	863	2	A48902	ATPase chain GyrA	1.95e+02	486	5	14.3	1272	2	S26180	neurofascin - chicken	1.95e+02
414	5	14.3	871	2	A46742	metabotropic glutamat	1.95e+02	487	5	14.3	1274	2	S59405	probable membrane pro	1.95e+02
415	5	14.3	874	3	T00491	capsid protein precu	1.95e+02	488	5	14.3	1279	2	E64709	type IIS restriction	1.95e+02
416	5	14.3	874	2	JQ0881	genome polyprotein -	1.95e+02	489	5	14.3	1298	1	EDBE75	immediate-early prote	1.95e+02
417	5	14.3	876	1	A57988	regulatory protein ar	1.95e+02	490	5	14.3	1308	2	E71622	probable membrane ass	1.95e+02
418	5	14.3	879	2	JH0562	metabotropic glutamat	1.95e+02	491	5	14.3	1311	2	A56390	mannosyl-glycoprotein	1.95e+02
419	5	14.3	890	2	H69877	calcium-transporting	1.95e+02	492	5	14.3	1317	2	S77517	DNA-directed RNA poly	1.95e+02
420	5	14.3	899	2	S17546	probable serine prote	1.95e+02	493	5	14.3	1353	2	JH0675	restrictin precursor	1.95e+02
421	5	14.3	900	2	F71436	hypothenical protein	1.95e+02	494	5	14.3	1355	1	VGBE11	149K glycoprotein - i	1.95e+02
422	5	14.3	905	2	S75035	H+-transporting ATPas	1.95e+02	495	5	14.3	1358	1	XOCHDH	xanthine dehydrogenas	1.95e+02
423	5	14.3	918	2	I58178	glutamate receptor -	1.95e+02	496	5	14.3	1379	2	S78371	DNA-directed RNA poly	1.95e+02
424	5	14.3	921	2	A48184	transcription initiat	1.95e+02	497	5	14.3	1396	2	A44453	translation initiatio	1.95e+02
425	5	14.3	921	2	A45183	TBP-associated factor	1.95e+02	498	5	14.3	1404	2	A36666	serate protein precu	1.95e+02
426	5	14.3	933	2	A31930	cytotactin - chicken	1.95e+02	499	5	14.3	1408	2	S16148	gene serrate protein	1.95e+02
427	5	14.3	935	2	S63261	SEC21 protein - yeast	1.95e+02	500	5	14.3	1417	2	T00561	hypothenical protein	1.95e+02
428	5	14.3	937	2	A56517	neocporin Nup98 - f	1.95e+02	501	5	14.3	1421	2	T02501	hypothenical protein	1.95e+02
429	5	14.3	941	2	T40772	hypothenical protein	1.95e+02	502	5	14.3	1433	2	S54587	CAT8 protein - yeast	1.95e+02
430	5	14.3	946	2	C65093	Iglutamate--ammonia-1	1.95e+02	503	5	14.3	1442	2	S57160	sulfite reductase hom	1.95e+02
431	5	14.3	948	2	T03225	probable regulatory p	1.95e+02	504	5	14.3	1447	2	S63669	UDPglucose--glycoprot	1.95e+02
432	5	14.3	949	2	S19808	glutamate receptor Gl	1.95e+02	505	5	14.3	1481	2	A47341	amylopullulanase prec	1.95e+02
433	5	14.3	952	2	S32954	hypothenical protein	1.95e+02	506	5	14.3	1481	2	S28669	amylopullulanase prec	1.95e+02
434	5	14.3	952	2	S64473	translation initiatio	1.95e+02	507	5	14.3	1487	2	S48719	phospholipase-A(2) re	1.95e+02
435	5	14.3	952	2	I50612	protein-tyrosine kina	1.95e+02	508	5	14.3	1516	2	T01055	hypothenical protein	1.95e+02
436	5	14.3	953	2	S12894	pyruvate,orthophospha	1.95e+02	509	5	14.3	1555	2	JT0959	polyprotein - potato	1.95e+02
437	5	14.3	953	2	S56650	pyruvate,orthophospha	1.95e+02	510	5	14.3	1557	2	T02859	probable membrane pro	1.95e+02
438	5	14.3	959	2	A55913	transcytosis-associat	1.95e+02	511	5	14.3	1609	2	T01797	hypothenical protein	1.95e+02
439	5	14.3	959	2	S35061	protovillin - slime m	1.95e+02	512	5	14.3	1615	2	B49502	protein-tyrosine-phos	1.95e+02
440	5	14.3	962	2	JT0669	helicase II-like prot	1.95e+02	513	5	14.3	1630	2	A53577	ascites sialoglycopro	1.95e+02
441	5	14.3	970	2	S48485	transcription activat	1.95e+02	514	5	14.3	1714	2	E71609	Ser/Thr protein kinas	1.95e+02
442	5	14.3	972	2	S77454	cation-transporting A	1.95e+02	515	5	14.3	1748	2	S42136	cnjB protein - Tetrah	1.95e+02
443	5	14.3	975	1	A28174	C1-tetrahydrofolate S	1.95e+02	516	5	14.3	1767	2	A49502	protein-tyrosine-phos	1.95e+02
444	5	14.3	982	2	S73741	p115 protein homolog	1.95e+02	517	5	14.3	1791	2	T02345	hypothenical protein	1.95e+02
445	5	14.3	991	2	T00058	hypothenical protein	1.95e+02	518	5	14.3	1797	2	F69195	cell surface glycopro	1.95e+02
446	5	14.3	991	2	T01372	hypothenical protein	1.95e+02	519	5	14.3	1810	2	A32230	tenascin precursor -	1.95e+02
447	5	14.3	991	2	T01377	hypothenical protein	1.95e+02	520	5	14.3	1819	2	A71928	cag island protein -	1.95e+02
448	5	14.3	999	2	S72267	Ca2+-transporting ATP	1.95e+02	521	5	14.3	1839	1	OYBYK	adenylate cyclase (EC	1.95e+02
449	5	14.3	999	2	A34307	Ca2+-transporting ATP	1.95e+02	522	5	14.3	1882	2	S73484	hypothetical protein	1.95e+02
450	5	14.3	1000	2	S44898	ZK1236.3 protein - Ca	1.95e+02	523	5	14.3	1937	2	G64585	cag pathogenicity isl	1.95e+02
451	5	14.3	1001	2	S74544	translation initiatio	1.95e+02	524	5	14.3	1938	2	A37361	probable integral mem	1.95e+02
452	5	14.3	1010	2	S45389	probable membrane pro	1.95e+02	525	5	14.3	2013	2	C71610	probable membrane ass	1.95e+02
453	5	14.3	1025	2	A43526	complement C3d/Epstei	1.95e+02	526	5	14.3	2014	2	S46622	probable membrane pro	1.95e+02
454	5	14.3	1027	2	I38604	p53-binding protein 1	1.95e+02	527	5	14.3	2145	2	JC4747	adenylate cyclase (EC	1.95e+02
455	5	14.3	1032	2	I38510	neuronal kinesin heav	1.95e+02	528	5	14.3	2164	1	GNV89	genome polyprotein -	1.95e+02
456	5	14.3	1033	2	S54506	probable membrane pro	1.95e+02	529	5	14.3	2285	2	A69911	hypothetical protein	1.95e+02
457	5	14.3	1040	2	S50617	BBB1 protein - yeast	1.95e+02	530	5	14.3	2388	2	JE0271	beta spectrin, beta-5	1.95e+02
458	5	14.3	1054	2	S54473	TPS3 protein - yeast	1.95e+02	531	5	14.3	2491	2	A57036	talin - slime mold (D	1.95e+02
459	5	14.3	1080	2	S30876	hypothenical protein	1.95e+02	532	5	14.3	2529	2	A56923	transcription factor	1.95e+02
460	5	14.3	1091	2	S57112	USJN1 protein - yeast	1.95e+02	533	5	14.3	2578	2	A56923	transcription factor	1.95e+02
461	5	14.3	1091	1	PL0009	complement C3d/Epstei	1.95e+02	534	5	14.3	2606	2	T03159	large tegument protei	1.95e+02

535	5	14.3	2843	1	RBHUP	adenomatous polyposis	1.95e+02	608	4	11.4	197	2	S22010	pyrimidine synthesis	3.99e+03
536	5	14.3	2845	2	I49505	adenomatous polyposis	1.95e+02	609	4	11.4	200	2	C49529	RNA-directed RNA poly	3.99e+03
537	5	14.3	3033	2	QJ1303	polyprotein - hepatitis	1.95e+02	610	4	11.4	207	2	D64601	phosphoserine phosphatase	3.99e+03
538	5	14.3	3169	2	T00296	toxin B - Escherichia	1.95e+02	611	4	11.4	214	2	S46476	cysteine proteinase (	3.99e+03
539	5	14.3	3566	2	A40701	tenascin-X precursor	1.95e+02	612	4	11.4	215	2	S23432	pyroglutamyl-peptidase	3.99e+03
540	5	14.3	3866	2	B48205	All-1 protein - GTE fo	1.95e+02	613	4	11.4	217	1	LCHUC	chorionomotropin A p	3.99e+03
541	5	14.3	3869	2	A48205	All-1 protein - GTE fo	1.95e+02	614	4	11.4	218	1	VCXUV	coat protein - cucum	3.99e+03
542	5	14.3	3924	2	S37431	ankyrin 2, neuronal 1	1.95e+02	615	4	11.4	219	2	JX0231	proteasome ring2 cha	3.99e+03
543	5	14.3	4391	2	A38096	perlecan precursor -	1.95e+02	616	4	11.4	221	2	I45939	MHC cell surface glyc	3.99e+03
544	5	14.3	4572	2	S67908	hypothetical polyprot	1.95e+02	617	4	11.4	222	2	F64757	membrane protein ykGH	3.99e+03
545	5	14.3	4910	2	S64942	probable membrane pro	1.95e+02	618	4	11.4	226	1	S37105	phosphoribosylglycine	3.99e+03
546	5	14.3	5037	2	A54161	ryanodine-binding pro	1.95e+02	619	4	11.4	231	2	JG6201	GSL protein homolog -	3.99e+03
547	5	14.3	6805	2	S20901	titin - rabbit (fragm	1.95e+02	620	4	11.4	236	2	F69912	hypothetical protein	3.99e+03
548	5	14.3	6839	2	S57242	titin - Caenorhabd	1.95e+02	621	4	11.4	237	2	D40595	hypothetical protein	3.99e+03
549	5	14.3	26926	1	I38344	titin, cardiac muscle	1.95e+02	622	4	11.4	238	1	LRRTB2	clathrin light chain	3.99e+03
550	4	11.4	39	2	S35325	protein kinase sggl0	3.99e+03	623	4	11.4	244	2	S70176	yeast protein - yersin	3.99e+03
551	4	11.4	56	2	S30243	Sry-type HMG box prot	3.99e+03	624	4	11.4	245	1	BIAG58	virB1 protein precurs	3.99e+03
552	4	11.4	87	2	QJ1264	hypothetical 10K prot	3.99e+03	625	4	11.4	246	2	G84147	hypothetical protein	3.99e+03
553	4	11.4	92	1	A31767	macrophage inflammato	3.99e+03	626	4	11.4	246	1	S01789	pyruvate formate-lyas	3.99e+03
554	4	11.4	95	1	TNLJ12	trans-activating tran	3.99e+03	627	4	11.4	247	2	C69954	hypothetical protein	3.99e+03
555	4	11.4	97	1	HVMS91	Ig heavy chain V regi	3.99e+03	628	4	11.4	249	2	C70436	conserved hypotheticala	3.99e+03
556	4	11.4	103	2	I64116	hypothetical protein	3.99e+03	629	4	11.4	252	2	S34747	anthranilate synthase	3.99e+03
557	4	11.4	106	2	S73492	hypothetical protein	3.99e+03	630	4	11.4	253	2	JQ2255	triose-phosphate isom	3.99e+03
558	4	11.4	115	2	A3925	proline-rich phosphop	3.99e+03	631	4	11.4	253	2	E84420	uroporphyrin-III C-me	3.99e+03
559	4	11.4	116	2	B69049	hypothetical protein	3.99e+03	632	4	11.4	255	2	E42725	nitrile hydratase 3-	3.99e+03
560	4	11.4	118	1	UTCAA	glycoprotein hormones	3.99e+03	633	4	11.4	255	2	A69939	conserved hypotheticala	3.99e+03
561	4	11.4	119	2	F27888	Ig heavy chain V-Ir	3.99e+03	634	4	11.4	256	2	H69201	hypothetical protein	3.99e+03
562	4	11.4	120	1	LHUV0	Ig lambda chain V-Ir	3.99e+03	635	4	11.4	258	1	VCOMCN	coat protein - cassav	3.99e+03
563	4	11.4	121	2	H27888	Ig heavy chain V regi	3.99e+03	636	4	11.4	260	2	JS0635	rRNA (adenine-N6-)me	3.99e+03
564	4	11.4	123	2	S29121	GTP-binding regulator	3.99e+03	637	4	11.4	261	2	B42882	motility protein (fla	3.99e+03
565	4	11.4	132	1	R3KT8	ribosomal protein S8	3.99e+03	638	4	11.4	261	1	WMMS	proliferating cell nu	3.99e+03
566	4	11.4	137	2	I71939	MHC class II I-A-beta	3.99e+03	639	4	11.4	270	2	I40724	hypothetical protein	3.99e+03
567	4	11.4	137	2	E64845	ydcV protein - Escher	3.99e+03	640	4	11.4	278	2	B40825	hypothetical protein	3.99e+03
568	4	11.4	138	2	S24081	envelope protein - he	3.99e+03	641	4	11.4	278	2	D59958	conserved hypotheticala	3.99e+03
569	4	11.4	138	2	C53291	probable B-ring cycla	3.99e+03	642	4	11.4	280	2	S76022	aspartate-semialdehyd	3.99e+03
570	4	11.4	140	2	S73680	MG337 homolog P02_0ff	3.99e+03	643	4	11.4	280	2	S55537	porin - fruit fly (Dr	3.99e+03
571	4	11.4	141	2	S64726	protein secretion pro	3.99e+03	644	4	11.4	284	2	S74256	homeotic nicotinate-n	3.99e+03
572	4	11.4	144	1	R3BY9E	ribosomal protein S19	3.99e+03	645	4	11.4	285	2	F70543	probable nicotinate-n	3.99e+03
573	4	11.4	144	2	H70063	hypothetical protein	3.99e+03	646	4	11.4	287	2	B65023	hypothetical protein	3.99e+03
574	4	11.4	145	2	S06307	T-cell receptor gamma	3.99e+03	647	4	11.4	288	2	S41006	hypothetical protein	3.99e+03
575	4	11.4	146	1	QVZ23	D2L protein - vaccini	3.99e+03	648	4	11.4	290	2	B64662	plasmid replication-p	3.99e+03
576	4	11.4	150	1	S49360	protein-tyrosine-phos	3.99e+03	649	4	11.4	290	2	H1852	hypothetical protein	3.99e+03
577	4	11.4	151	2	S23215	hypothetical protein	3.99e+03	650	4	11.4	291	2	I48602	insulin-like growth f	3.99e+03
578	4	11.4	151	2	QJ1794	SalG2R protein - vacc	3.99e+03	651	4	11.4	296	2	H70585	hypothetical protein	3.99e+03
579	4	11.4	157	2	S47575	EGP1 protein - yeast	3.99e+03	652	4	11.4	299	2	A35272	osteoinductive factor	3.99e+03
580	4	11.4	158	2	B64117	transcription elongat	3.99e+03	653	4	11.4	301	2	I39754	nodp protein - Azosp	3.99e+03
581	4	11.4	159	2	B29879	vasotocin / neurophys	3.99e+03	654	4	11.4	304	2	F36808	hypothetical protein	3.99e+03
582	4	11.4	159	2	A48428	homeotic protein Hox	3.99e+03	655	4	11.4	305	2	I39376	translation initiatio	3.99e+03
583	4	11.4	163	1	CFVCA	C-Phycocyanin alpha c	3.99e+03	656	4	11.4	307	2	A32208	synaptophysin - bovin	3.99e+03
584	4	11.4	164	1	GCCHH	glycine cleavage syst	3.99e+03	657	4	11.4	308	2	G64781	probable transcriptio	3.99e+03
585	4	11.4	166	2	G69911	hypothetical protein	3.99e+03	658	4	11.4	311	2	A65010	d-serine deaminase ac	3.99e+03
586	4	11.4	166	1	WNADP5	early ELB 20K protein	3.99e+03	659	4	11.4	315	2	F70203	xylose operon regulat	3.99e+03
587	4	11.4	167	2	S21359	keratin, type I, cyto	3.99e+03	660	4	11.4	316	1	QBEN3	gene 34 protein - sai	3.99e+03
588	4	11.4	168	2	A03864	hypothetical protein	3.99e+03	661	4	11.4	317	2	S35282	eaa protein - phage P	3.99e+03
589	4	11.4	170	2	C70102	hypothetical protein	3.99e+03	662	4	11.4	319	1	ERADF3	fiber protein - human	3.99e+03
590	4	11.4	172	2	B29149	proline-rich protein	3.99e+03	663	4	11.4	319	1	RWRBB	T-cell receptor beta	3.99e+03
591	4	11.4	173	1	GCCHH	glycine cleavage syst	3.99e+03	664	4	11.4	320	1	DCLLHP	histidine decarboxyla	3.99e+03
592	4	11.4	173	2	A53729	myosin light chain, c	3.99e+03	665	4	11.4	321	2	H71729	hypothetical protein	3.99e+03
593	4	11.4	175	2	S70915	major oleosin nap-II	3.99e+03	666	4	11.4	323	2	F70141	probable oligopeptid	3.99e+03
594	4	11.4	177	2	C40428	nonspecific cross-rea	3.99e+03	667	4	11.4	323	2	A26702	arginase (EC 3.5.3.1)	3.99e+03
595	4	11.4	178	2	S14505	hypothetical protein	3.99e+03	668	4	11.4	326	2	S23387	protein kinase (EC 2.	3.99e+03
596	4	11.4	178	2	S53018	hypothetical protein	3.99e+03	669	4	11.4	330	2	H70728	probable atp-binding	3.99e+03
597	4	11.4	178	2	G70134	flagellar protein (fl	3.99e+03	670	4	11.4	330	2	S74255	homeotic protein six3	3.99e+03
598	4	11.4	179	2	A05248	170K protein - tobacc	3.99e+03	671	4	11.4	331	2	S01964	readthrough protein -	3.99e+03
599	4	11.4	179	2	E36786	hypothetical protein	3.99e+03	672	4	11.4	332	2	B70965	probable integrase	3.99e+03
600	4	11.4	186	2	F70230	conserved hypotheticala	3.99e+03	673	4	11.4	334	2	G71194	probable glyceraldehy	3.99e+03
601	4	11.4	192	2	A70331	peptidyl-trNA hydrola	3.99e+03	674	4	11.4	335	1	DELBC	D-2-hydroxyisocapro	3.99e+03
602	4	11.4	193	1	JWBPV	DNA-invertase - phage	3.99e+03	675	4	11.4	337	2	C70473	phosphate-binding per	3.99e+03
603	4	11.4	193	2	A69155	hypothetical protein	3.99e+03	676	4	11.4	338	2	S12731	L-serine dehydratase	3.99e+03
604	4	11.4	195	2	S53112	delta antigen - hepat	3.99e+03	677	4	11.4	338	2	E64417	transcription initiat	3.99e+03
605	4	11.4	195	2	S00157	superoxide dismutase	3.99e+03	678	4	11.4	338	2	S16116	MHC HLA-B27-HS - huma	3.99e+03
606	4	11.4	195	2	S22475	oleosin BN-III - rape	3.99e+03	679	4	11.4	340	2	H70950	hypothetical protein	3.99e+03
607	4	11.4	197	2	S09945	hypothetical protein	3.99e+03	680	4	11.4	341	2	JH0769	glycerolaldehyde-3-phos	3.99e+03

681	4	11.4	343	2	C70418	alcohol dehydrogenase	3.99e+03	1	S51031	acid phosphatase (EC	3.99e+03
682	4	11.4	343	2	C69966	hypothetical protein	3.99e+03	432	S18407	acrosin (EC 3.4.21.10	3.99e+03
683	4	11.4	347	2	A71903	d-alanine-d-alanine	3.99e+03	437	A31142	gelsolin, ovarian - A	3.99e+03
684	4	11.4	347	2	S40082	hypothetical protein	3.99e+03	438	H64713	Na+/H+-exchanging pro	3.99e+03
685	4	11.4	348	2	S76587	hypothetical protein	3.99e+03	438	S69965	TyA protein - yeast (	3.99e+03
686	4	11.4	348	2	A23665	opsin - mouse	3.99e+03	439	I60916	HNF-3/forkhead homolo	3.99e+03
687	4	11.4	350	2	H70929	probable PPE protein	3.99e+03	441	C69834	sensory transduction	3.99e+03
688	4	11.4	353	2	S69730	hypothetical protein	3.99e+03	443	C70780	probable PPE protein	3.99e+03
689	4	11.4	354	2	I80168	class I histocompatib	3.99e+03	443	A46248	dihydroorotate oxidas	3.99e+03
690	4	11.4	358	2	H36891	transfer complex prot	3.99e+03	447	A38561	alpha-1,3-mannosyl-gl	3.99e+03
691	4	11.4	358	2	S03538	class I histocompatib	3.99e+03	447	S18332	NADH peroxidase (EC 1	3.99e+03
692	4	11.4	359	2	A48073	protein-tyrosine kina	3.99e+03	448	B42022	transcription factor	3.99e+03
693	4	11.4	359	1	HLH084	MHC class I histocomp	3.99e+03	449	S73808	glycine--trNA ligase	3.99e+03
694	4	11.4	359	2	H64547	peptide methionine su	3.99e+03	450	I50484	vimentin beta - goldf	3.99e+03
695	4	11.4	359	2	S71963	GTP-binding protein a	3.99e+03	451	S30436	protein kinase PCRAIR	3.99e+03
696	4	11.4	361	1	I64139	translating releasing	3.99e+03	452	JC4100	hydroxyindole O-methy	3.99e+03
697	4	11.4	361	1	DENTN2	NADH dehydrogenase (u	3.99e+03	452	S16522	mitosis-specific cycl	3.99e+03
698	4	11.4	362	2	JH0288	class I histocompatib	3.99e+03	454	D69066	amidease - Methanobact	3.99e+03
699	4	11.4	362	2	S25369	3-isopropylmalate deh	3.99e+03	454	F70564	probable serine prote	3.99e+03
700	4	11.4	362	1	HLH088	MHC class I histocomp	3.99e+03	457	S43137	D-CD36 protein - frui	3.99e+03
701	4	11.4	362	2	I37485	human lymphocyte anti	3.99e+03	459	A34791	interleukin-7 recepto	3.99e+03
702	4	11.4	363	2	S30386	mannan endo-1,4-beta-	3.99e+03	462	A24201	macrophage elastase (	3.99e+03
703	4	11.4	364	2	F36470	Wnt-6 protein - mouse	3.99e+03	462	D70936	probable aroG - Mycob	3.99e+03
704	4	11.4	364	2	D35997	MHC class I histocomp	3.99e+03	465	1 XIECMB	5-methylcytosine-spec	3.99e+03
705	4	11.4	364	2	I39048	alpha (1,3) fucosyltr	3.99e+03	466	S39494	glutathione reductase	3.99e+03
706	4	11.4	365	2	S17372	sigma 3 protein - reo	3.99e+03	466	A23685	interstitial collagen	3.99e+03
707	4	11.4	371	1	R6BYW7	ribosomal protein YML	3.99e+03	466	S62330	beta-fructofuranosida	3.99e+03
708	4	11.4	372	2	A25148	thymus leukemia antiq	3.99e+03	466	S62332	beta-fructofuranosida	3.99e+03
709	4	11.4	373	2	A41948	transforming growth f	3.99e+03	467	S36559	L2 protein - human pa	3.99e+03
710	4	11.4	374	1	A53142	alcohol dehydrogenase	3.99e+03	469	I37451	HBFG-2 (HFK-2) protei	3.99e+03
711	4	11.4	374	1	A56436	alcohol dehydrogenase	3.99e+03	471	S68463	protein kinase ATPK19	3.99e+03
712	4	11.4	375	1	DMSAA	alcohol dehydrogenase	3.99e+03	471	A53711	collagenase 3 (EC 3.4	3.99e+03
713	4	11.4	375	1	S66272	alcohol dehydrogenase	3.99e+03	472	S29243	interstitial collagen	3.99e+03
714	4	11.4	377	2	S27671	rip protein - Shigell	3.99e+03	472	1 KRHUE	keratin, 50K type I c	3.99e+03
715	4	11.4	378	2	B4581	probable potassium ch	3.99e+03	473	C64504	hypothetical protein	3.99e+03
716	4	11.4	379	1	S04571	alcohol dehydrogenase	3.99e+03	474	2 D43943	essential for lactoco	3.99e+03
717	4	11.4	379	1	DEILSP	alcohol dehydrogenase	3.99e+03	475	C64075	aspartate ammonia-ly	3.99e+03
718	4	11.4	379	1	S57819	alcohol dehydrogenase	3.99e+03	476	H71936	proteinase DO - Helic	3.99e+03
719	4	11.4	380	1	S53307	alcohol dehydrogenase	3.99e+03	478	J50764	aspartate ammonia-ly	3.99e+03
720	4	11.4	382	2	A28067	lysosomal membrane gl	3.99e+03	478	A54040	adenosylhomocysteinas	3.99e+03
721	4	11.4	382	1	DEPJA1	alcohol dehydrogenase	3.99e+03	485	C71400	probable adenosylhomo	3.99e+03
722	4	11.4	384	2	A25132	MHC class I histocomp	3.99e+03	485	S38379	adenosylhomocysteinas	3.99e+03
723	4	11.4	385	2	S49752	homeotic protein YOX1	3.99e+03	486	F69762	transporter homolog y	3.99e+03
724	4	11.4	385	2	C71699	murg protein (murg) R	3.99e+03	487	S71170	calcium-dependent pro	3.99e+03
725	4	11.4	387	2	A53586	albumin-binding prote	3.99e+03	488	S18156	globulin 1 - eastern	3.99e+03
726	4	11.4	389	2	S75506	large surface antigen	3.99e+03	488	1 QXASBI	mENA maturase bil - E	3.99e+03
727	4	11.4	390	1	A39249	neutrophil cytosol fa	3.99e+03	492	S28003	trypanothione reducta	3.99e+03
728	4	11.4	391	2	D70922	probable PPE protein	3.99e+03	495	S31493	env polypeptide - hum	3.99e+03
729	4	11.4	392	1	A24608	gastricsin (EC 3.4.23	3.99e+03	496	1 JQ1882	envelope glycoprotein	3.99e+03
730	4	11.4	393	2	J00410	methionine adenosyltr	3.99e+03	498	PC6300	synaptotagmin X - rat	3.99e+03
731	4	11.4	394	2	S26431	intermediate filament	3.99e+03	498	1 W2WL8	E2 protein - human pa	3.99e+03
732	4	11.4	395	1	S38821	homeotic protein lim-	3.99e+03	499	B64069	fructose phosphotrans	3.99e+03
733	4	11.4	396	2	H70741	probable PPE protein	3.99e+03	500	1 ISEBAB	L-arabinose isomerase	3.99e+03
734	4	11.4	396	2	PS0075	myosin heavy chain, c	3.99e+03	502	2 A23547	keratin, type II cyto	3.99e+03
735	4	11.4	399	1	A39697	maize myb-related pro	3.99e+03	502	JN0113	nicotinic acetylcholi	3.99e+03
736	4	11.4	402	2	D70186	xylose operon regulat	3.99e+03	502	S69331	natriuretic peptide r	3.99e+03
737	4	11.4	405	2	D64522	carboxymorspermidine	3.99e+03	503	S59698	HST1 protein - yeast	3.99e+03
738	4	11.4	407	2	A44374	3-carboxy-cis,cis-muc	3.99e+03	504	S33194	phase-1 flagellin - S	3.99e+03
739	4	11.4	408	2	S48813	probable chain length	3.99e+03	505	S15193	sucrose porin scry -	3.99e+03
740	4	11.4	408	2	S11320	serine proteinase inh	3.99e+03	505	1 TVHUHC	protein-tyrosine kina	3.99e+03
741	4	11.4	409	2	D64216	hypothetical protein	3.99e+03	506	JN0615	steroid 11beta-monoox	3.99e+03
742	4	11.4	409	1	VCVW2S	env polypeptide (vers	3.99e+03	507	S33185	phase-1 flagellin - S	3.99e+03
743	4	11.4	410	2	PN0008	phosphoglycerate kina	3.99e+03	507	A30828	steroid 17alpha-monoo	3.99e+03
744	4	11.4	411	2	B50008	phosphoglycerate kina	3.99e+03	508	F71674	methionyl-tRNA synth	3.99e+03
745	4	11.4	411	1	HFSNFA	fibrolase (EC 3.4.24.	3.99e+03	510	1 DERNZ2	NADH dehydrogenase (u	3.99e+03
746	4	11.4	414	2	A29835	SalI protein - Escher	3.99e+03	513	2 S08381	cholesterol monooxyl	3.99e+03
747	4	11.4	415	2	A34170	acrosin (EC 3.4.21.10	3.99e+03	514	S32197	keratin, 58K type II,	3.99e+03
748	4	11.4	422	2	S49012	yolk protein 3 - blue	3.99e+03	514	1 A38668	IMP dehydrogenase (EC	3.99e+03
749	4	11.4	423	2	S60939	GTP-binding protein b	3.99e+03	519	2 S69989	unspecific monooxygen	3.99e+03
750	4	11.4	426	2	S70396	zona pellucida glycop	3.99e+03	520	1 FOLJGL	gaq polypeptide - gib	3.99e+03
751	4	11.4	426	2	B55334	pregnancy-specific be	3.99e+03	521	S15070	amidase (EC 3.5.1.4)	3.99e+03
752	4	11.4	429	2	A25145	keratin, 47K type I c	3.99e+03	522	2 G02533	occludin - human	3.99e+03
753	4	11.4	430	2	S35960	aspartate transaminas	3.99e+03	524	1 VHBP20	capsid protein 20 - p	3.99e+03





```
973 4 11.4 1453 2 S21626 collagen alpha 1(I) c 3.99e+03
974 4 11.4 1456 2 A36563 mannose receptor prec 3.99e+03
975 4 11.4 1496 2 I51170 carbamyl phosphate sy 3.99e+03
976 4 11.4 1639 1 MNFFB2 laminin gamma-1 chain 3.99e+03
977 4 11.4 1823 2 S28974 vitellogenin precursor 3.99e+03
978 4 11.4 1902 2 B44858 serine proteinase (EC 3.99e+03
979 4 11.4 1935 1 S06006 myosin beta heavy cha 3.99e+03
980 4 11.4 1938 1 S06005 myosin alpha heavy ch 3.99e+03
981 4 11.4 1939 1 A46762 myosin alpha heavy ch 3.99e+03
982 4 11.4 2185 1 GNNYB3 genome polyprotein - 3.99e+03
983 4 11.4 2206 1 GNNY4P genome polyprotein - 3.99e+03
984 4 11.4 2206 1 GNNY27 genome polyprotein - 3.99e+03
985 4 11.4 2209 1 GNNY3P genome polyprotein - 3.99e+03
986 4 11.4 2214 1 Q2BYU2 pyrimidine synthesis 3.99e+03
987 4 11.4 2242 2 A57541 pyrimidine synthesis 3.99e+03
988 4 11.4 2248 2 A35938 profilaggrin - human 3.99e+03
989 4 11.4 2291 2 S11238 polymerase - Berne vi 3.99e+03
990 4 11.4 2385 2 A32491 myosin heavy chain 1, 3.99e+03
991 4 11.4 2477 2 S14428 fibronectin precursor 3.99e+03
992 4 11.4 2512 1 X1CHFA fatty-acid synthase ( 3.99e+03
993 4 11.4 2663 1 S28261 centromere protein E 3.99e+03
994 4 11.4 2688 2 I49477 alpha-A-crystallin-bi 3.99e+03
995 4 11.4 2893 2 A84536 toxin-like outer memb 3.99e+03
996 4 11.4 3011 1 GNMVCH genome polyprotein - 3.99e+03
997 4 11.4 3388 1 GNMVDP genome polyprotein - 3.99e+03
998 4 11.4 3391 1 GNMV16 genome polyprotein - 3.99e+03
999 4 11.4 3432 1 GNMVUS genome polyprotein - 3.99e+03
1000 4 11.4 4543 1 A53102 alpha-2-macroglobulin 3.99e+03
```

## ALIGNMENTS

```
RESULT 1
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#submissions
#description
#accession
#status
#molecule_type
#residues
#cross-references
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT 2
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#molecule_type
#residues
#cross-references
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT 4
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#molecule_type
#residues
#cross-references
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT 4
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#molecule_type
#residues
#cross-references
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
```

```
#title
#accession
#status
#molecule_type
#residues
#cross-references
#experimental_source
GENETICS
#map_position
#note
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT 3
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#accession
#molecule_type
#residues
CLASSIFICATION
KEYWORDS
FEATURE
SUMMARY
Query Match
Best Local Similarity
Matches
Db
Qy
RESULT 4
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title
#molecule_type
#residues
CLASSIFICATION
```

```
calmodulin repeat homology
calcium binding; EF hand
#domain calmodulin repeat homology #label EF4
#length 183 #molecular-weight 21118 #checksum 4156
Query Match 17.1%; Score 6; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69 INKDDV 74
QY 16 INKDDV 21

RESULT 5
ENTRY A71230 #type complete
TITLE hypothetical protein PH0099 - Pyrococcus horikoshii
ORGANISM #formal_name Pyrococcus horikoshii
DATE 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change
ACCESSIONS A71230
REFERENCE #authors Kavarabavasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.;
Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
Horikoshi, K.; Masuchi, Y.; Shizuza, H.; Kikuchi, H.
#journal DNA Res. (1998) 5:55-76
#title Complete sequence and gene organization of the genome of a
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii
OT3.
#cross-references MUID:98344137
#accession A71230
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-184 #label KAW
#cross-references GB:AP000001; NID:g3236128; PID:d1030111; PID:g3256485
#experimental_source strain OT3
#note this accession replaces an interim accession for a
sequence replaced by GenBank

GENETICS
#gene PH0099
SUMMARY #length 184 #molecular-weight 21095 #checksum 6781
Query Match 17.1%; Score 6; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 IENDEI 99
QY 2 IENDEI 7

RESULT 6
ENTRY S13183 #type complete
TITLE calcium-binding protein SCP II - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
lancelet
DATE 02-Dec-1993 #sequence_revision 08-Sep-1995 #text_change
ACCESSIONS S13183
REFERENCE #authors Takagi, T.; Cox, J.A.
#journal Eur. J. Biochem. (1990) 192:387-399
#title Amino acid sequences of four isoforms of amphioxus
sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13183
#status preliminary
#molecule_type protein

1-185 #label TAK
#superfamily unassigned calmodulin-related proteins;
calmodulin repeat homology
alternative splicing; calcium binding; EF hand
#domain calmodulin repeat homology #label EF1\
#domain calmodulin repeat homology #label EF2\
#domain calmodulin repeat homology #label EF3\
#domain calmodulin repeat homology #label EF4
#length 185 #molecular-weight 21286 #checksum 6749
Query Match 17.1%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 INKDDV 76
QY 16 INKDDV 21

RESULT 7
ENTRY S13184 #type complete
TITLE calcium-binding protein SCP III - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
lancelet
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
ACCESSIONS S13184
REFERENCE #authors Takagi, T.; Cox, J.A.
#journal Eur. J. Biochem. (1990) 192:387-399
#title Amino acid sequences of four isoforms of amphioxus
sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13184
#status preliminary
#molecule_type protein

1-185 #label TAK
#superfamily unassigned calmodulin-related proteins;
calmodulin repeat homology
#domain calmodulin repeat homology #label EF1\
#domain calmodulin repeat homology #label EF2\
#domain calmodulin repeat homology #label EF3\
#domain calmodulin repeat homology #label EF4
#length 185 #molecular-weight 21386 #checksum 6655
Query Match 17.1%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 INKDDV 76
QY 16 INKDDV 21

RESULT 8
ENTRY S13185 #type complete
TITLE calcium-binding protein SCP IV - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
lancelet
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
ACCESSIONS S13185
REFERENCE #authors Takagi, T.; Cox, J.A.
#journal Eur. J. Biochem. (1990) 192:387-399
#title Amino acid sequences of four isoforms of amphioxus
sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13185
#status preliminary
#molecule_type protein
```

```

##residues      1-185 #label TAK
##note          the sequence from Table 4 is inconsistent with that from
                Fig. 2 in lacking 43-Lys, 98-Asp, 100-Pro, and 145-Tyr
CLASSIFICATION #superfamily unassigned calmodulin-related proteins;
                calmodulin repeat homology
FEATURE
6-38           #domain calmodulin repeat homology #label EF1\
57-89           #domain calmodulin repeat homology #label EF2\
102-134         #domain calmodulin repeat homology #label EF3\
136-168         #domain calmodulin repeat homology #label EF4\
SUMMARY        #length 185 #molecular-weight 21417 #checksum 7125
Query Match    17.1%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 INKDDV 76
QY 16 INKDDV 21
|||||

RESULT 9
ENTRY   S13182 #type complete
TITLE   calcium-binding protein SCP I - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
          lancelet
DATE    19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
          07-Aug-1998
ACCESSIONS S13182
REFERENCE  Takagi, T.; Cox, J.A.
          Eur. J. Biochem. (1990) 192:387-399
#journal
#title     Amino acid sequences of four isoforms of amphioxus
          sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13182
#status preliminary
#molecule_type protein
#residues 1-185 #label TAK
CLASSIFICATION #superfamily unassigned calmodulin-related proteins;
                calmodulin repeat homology
FEATURE
6-38           #domain calmodulin repeat homology #label EF1\
57-89           #domain calmodulin repeat homology #label EF2\
102-134         #domain calmodulin repeat homology #label EF3\
136-168         #domain calmodulin repeat homology #label EF4\
SUMMARY        #length 185 #molecular-weight 21418 #checksum 6895
Query Match    17.1%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 INKDDV 76
QY 16 INKDDV 21
|||||

RESULT 10
ENTRY   E70079 #type complete
TITLE   conserved hypothetical protein yxjg - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE    05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
          24-Sep-1998
ACCESSIONS E70079
REFERENCE  A69580
          Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
          Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
          Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
          A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
          Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
          Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;
          Daniel, R.A.; Denizot, F.; Devigne, K.M.; Duesterhoeft, A.;
          Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

```

```

Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita, M.;
Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.;
Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppi, G.;
Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.;
Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.;
Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.;
Lapidus, A.; Lardinois, S.; Lauber, J.; Lazarevic, V.; Lee, S.M.;
Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.;
Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.;
Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.;
Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.;
Porwolik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.;
Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.;
Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.;
Sato, T.; Scanlon, E.; Schleich, S.; Schroeter, R.; Scoffone, F.;
Sekiguchi, J.; Sekowka, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.;
Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.;
Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.;
Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.;
Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.;
Wedler, H.; Weitzenecker, T.; Winters, P.; Wipat, A.; Yamamoto, H.;
Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.;
Zumstein, E.; Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references MUID:98044033
#accession E70079
#status preliminary; nucleic acid sequence not shown;
          translation not shown
#molecule_type DNA
#residues 1-334 #label KUN
#cross-references GB:299123; GB:AL009126; NID:g2636240; PID:e1186395;
          PID:g2636431
#experimental_source strain 168
GENETICS
#gene yxjg
SUMMARY #length 334 #molecular-weight 38017 #checksum 7798
Query Match 17.1%; Score 6; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 54 IENDEI 59
QY 2 IENDEI 7
|||||

RESULT 11
ENTRY   A25004 #type fragment
TITLE   keratin, 53k type II cytoskeletal - bovine (fragment)
ALTERNATE_NAMES cytokeatin A (number 8)
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE    25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change
          18-Jun-1993
ACCESSIONS A25004
REFERENCE  Magin, T.M.; Jorcano, J.L.; Franke, W.W.
          Differentiation (1986) 30:254-264
          Cytokeatin expression in simple epithelia. II. CDNA cloning
          and sequence characteristics of bovine cytokeatin A (no.
          8).
#cross-references MUID:86193259
#accession A25004
#molecule_type mRNA
#residues 1-370 #label MAG
CLASSIFICATION #superfamily cytoskeletal keratin
KEYWORDS coiled coil
SUMMARY #length 370 #checksum 5736

```

```

Query Match      17.1% Score 6; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 340 FSRSS 345
      |||||
Qy 28 FSRSS 33

RESULT 12
ENTRY 12
TITLE 12
ORGANISM 12
DATE 12-Aug-1998
ACCESSIONS F69031
REFERENCE F69031
#authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
          Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
          Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
          Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.;
          Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; Caruso,
          A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
          McDougall, S.; Shimer, G.; Goyal, A.; Pietrowski, S.;
          Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
          J.; Reeve, J.N.
#journal J. Bacteriol. (1997) 179:7135-7155
#title Complete genome sequence of Methanobacterium
        thermoautotrophicum Delta H: functional analysis and
        comparative genomics.
#cross-references MUID:98037514
#accession F69031
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-423 #label MTH
##cross-references GB:A6000890; GB:AE000666; NID:g2622331; PID:g2622343
##experimental_source strain Delta H

GENETICS
#gene MTH1232
#start_codon GTG
#classification #superfamily homoserine dehydrogenase; homoserine
        dehydrogenase homology
KEYWORDS NADP: oxidoreductase; threonine biosynthesis
SUMMARY #length 423 #molecular-weight 45913 #checksum 9156

Query Match      17.1% Score 6; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 338 FSRSS 343
      |||||
Qy 28 FSRSS 33

RESULT 13
ENTRY 13
TITLE 13
ORGANISM 13
DATE 14-Aug-1998
ACCESSIONS B71153
REFERENCE B71153
#authors Kavarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.;
          Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
          Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
          Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
          Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.; Oguchi,
          A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
          Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.
#journal DNA Res. (1998) 5:55-76

#title Complete sequence and gene organization of the genome of a
        hyper-thermophilic archaeobacterium, Pyrococcus horikoshii
        OT3.
#cross-references MUID:98344137
#accession B71153
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-440 #label KAW
##cross-references GB:AF000002; NID:g3236129; PID:d1030454; PID:g3256828
##experimental_source strain OT3
#note this accession replaces an interim accession for a
        sequence replaced by GenBank

GENETICS
#gene PH0425
SUMMARY #length 440 #molecular-weight 50609 #checksum 2190

Query Match      17.1% Score 6; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 366 INKDDV 371
      |||||
Qy 16 INKDDV 21

RESULT 14
ENTRY 14
TITLE 14
ORGANISM 14
DATE 06-Jul-1990
ACCESSIONS A34720
REFERENCE A34720
#authors Yamamoto, R.; Kao, L.C.; McKnight, C.E.; Strauss III, J.F.
#journal Mol. Endocrinol. (1990) 4:370-374
#title Cloning and sequence of cDNA for human placental cyokeratin
        8. Regulation of the mRNA in trophoblastic cells by CAMP.
#cross-references MUID:90258929
#accession A34720
#status preliminary
#molecule_type mRNA
#residues 1-483 #label YAM
##cross-references GB:M34225
REFERENCE S06888
#authors Leube, R.E.; Bosch, F.X.; Romano, V.; Zimbelmann, R.;
        Hoefler, H.; Franke, W.W.
#journal Differentiation (1986) 33:69-85
#title Cyokeratin expression in simple epithelia.
#cross-references MUID:87134779
#accession S06888
#molecule_type mRNA
#residues 205-483 #label LEU
##cross-references EMBL:X12882
REFERENCE S12479
#authors Franke, W.W.
#submission submitted to the EMBL Data Library, September 1988
#accession S12479
#molecule_type mRNA
#residues 205-309 'I', 311-483 #label FRA
##cross-references EMBL:X12882; NID:g30312; PID:g30313
REFERENCE I37982
#authors Waseem, A.; Alexander, C.M.; Steel, J.B.; Lane, E.B.
#journal New Biol. (1990) 2:464-478
#title Embryonic simple epithelial keratins 8 and 18: chromosomal
        location emphasizes difference from other keratin pairs.
#cross-references MUID:91145351
#accession I37982
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-76 'V', 78-416, 'S', 418-428, 'G', 430-431, 'S', 433-483
        #label RES
##cross-references EMBL:X74929; NID:g400415; PID:g400416

```

```
#accession      I37983
##status        preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues      151-201.'VP'.202-271 ##label RE2
##cross-references EMBL:X74981; NID:g400417; PID:g400418
REFERENCE
JS0487
#authors        Krauss, S.; Franke, W.W.
#journal         Gene (1990) 86:241-249
#title           Organization and sequence of the human gene encoding
                 cytokeratin 8.
#cross-references MUID:90215304
#accession      JS0487
##molecule_type DNA
##residues      1-76.'V'.78-428.'G'. 'SOA'.434-483 ##label KRA
##cross-references GB:M34482; NID:g181572; PID:g181573
REFERENCE
A31942
#authors        Kulesh, D.A.; Cecena, G.; Darmon, Y.M.; Vasseur, M.; Oshima,
                 R.G.
#journal         Mol. Cell. Biol. (1989) 9:1553-1565
#title           Posttranslational regulation of keratins: degradation of
                 mouse and human keratins 18 and 8.
#cross-references MUID:89261783
#accession      A31942
##status        preliminary
##molecule_type mRNA
##residues      1-76.'V'.78-231 ##label KUL
##cross-references GB:M26512; NID:g17796; PID:g553163
COMMENT          This protein is one of the type II cytokeratins.
GENETICS
#gene           GDB:KRT8
##cross-references GDB:118830; OMIM:148060
#map_position   12p13.2-12q24.1
#introns        108/3; 178/2; 198/3; 230/3; 327/3; 401/2; 421/1
CLASSIFICATION #superfamily cytoskeletal keratin
KEYWORDS        coiled coil; intermediate filament
SUMMARY         #length 483 #molecular-weight 53748 #checksum 3427

Query Match      17.1%; Score 6; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 5.36e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 452 FSRSS 457
|||||
QY 28 FSRSS 33

RESULT 15
ENTRY
TITLE          #type complete
                Probable thymidine phosphorylase (EC 2.4.2.4) 2 -
                Archaeoglobus fulgidus
ORGANISM        #formal_name Archaeoglobus fulgidus
DATE            23-Oct-1998 #sequence_revision 23-Oct-1998 #text_change
                23-Oct-1998
ACCESSIONS
REFERENCE       A69250
#authors        Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
                K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.;
                Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham,
                D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
                Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
                Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.;
                Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.;
                Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
                J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
                T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.;
                D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.;
                Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
                C.R.; Venter, J.C.
#journal         Nature (1997) 390:364-370
#title           The complete genome sequence of the hyperthermophilic,
                 sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references MUID:98049343
#accession      E69417
```

\*\*\*\*\*  
W I S K E H  
\*\*\*\*\* (TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:05:23 1999; MasPar time 3.51 Seconds  
Tabular output not generated. 281.812 Million cell updates/sec

Title: >US-09-049-696-48  
Description: (1-35) from US09049696.pep  
Perfect Score: 35  
Sequence: 1 WIENDEIQWNPPEINRDKVDQHKQVCFSTSSG 35

Scoring table: TABLE uniprotatable  
Gap 60

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 2.750; Variance 0.395; scale 6.956

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	17.1	185	1	SCP2_BRALA SARCOPLASMIC CALCIUM-B	2.56e+00
2	6	17.1	185	1	SCP1_BRALA SARCOPLASMIC CALCIUM-B	2.56e+00
3	6	17.1	334	1	YXJG_BACSU HYPOTHETICAL 38.0 KD P	2.56e+00
4	6	17.1	370	1	K2C8_BOVIN KERATIN, TYPE II CYTOS	2.56e+00
5	6	17.1	482	1	K2C8_HUMAN KERATIN, TYPE II CYTOS	2.56e+00
6	6	17.1	670	1	SR72_CANPA SIGNAL RECOGNITION PAR	2.56e+00
7	6	17.1	670	1	SR72_HUMAN SIGNAL RECOGNITION PAR	2.56e+00
8	6	17.1	877	1	MGR3_HUMAN METABOTROPIC GLUTAMATE	2.56e+00
9	6	17.1	903	1	ECLC_BOVIN EPIHELIAL CHLORIDE CH	2.56e+00
10	6	17.1	916	1	HMDH_DROME 3-HYDROXY-3-METHYLGLUT	2.56e+00
11	6	17.1	1681	1	CLH_CAEEL PROBABLE CLATHRIN HEAV	2.56e+00
12	6	17.1	1967	1	RPO_PVMR RNA REPLICATION PROTEI	2.56e+00
13	6	17.1	2205	1	POLN_RUBVT NONSTRUCTURAL POLYPROT	2.56e+00
14	6	17.1	2703	1	NOTC_DROME NEUROGENIC LOCUS NOTCH	2.56e+00
15	5	14.3	39	1	ABAE_BOMPA ABACIN.	1.14e+02
16	5	14.3	60	1	MERC_PSEAE HYPOTHETICAL 7.0 KD PR	1.14e+02
17	5	14.3	60	1	Y02B_BPT4 HYPOTHETICAL MERCURIC	1.14e+02
18	5	14.3	60	1	MERC_SHIFL DNA-DIRECTED RNA POLYM	1.14e+02
19	5	14.3	63	1	RPCX_SCHPO DNA-DIRECTED RNA POLYM	1.14e+02
20	5	14.3	81	1	YIDD_PSEPU HYPOTHETICAL 9.2 KD PR	1.14e+02
21	5	14.3	105	1	REV_HPV18 E7 PROTEIN.	1.14e+02
22	5	14.3	106	1	REV_HV18 REV PROTEIN (ANTI-REPR	1.14e+02
23	5	14.3	107	1	Y453_METJA HYPOTHETICAL PROTEIN M	1.14e+02

1	107	14.3	5	14.3	107	1	FERREDOXIN PRECURSOR.	1.14e+02
1	119	14.3	5	14.3	119	1	YVAA_VACCC HYPOTHETICAL 14.1 KD P	1.14e+02
1	120	14.3	5	14.3	120	1	NEUB_XENLA NEUROMEDIN B PRECURSOR	1.14e+02
1	130	14.3	5	14.3	130	1	TAT_HV2RO TAT PROTEIN (TRANSACTI	1.14e+02
1	131	14.3	5	14.3	131	1	ATPE_GUIHT ATP SYNTHASE EPSILON C	1.14e+02
1	133	14.3	5	14.3	133	1	XTP1_YEAST HYPOTHETICAL 15.2 KD P	1.14e+02
1	138	14.3	5	14.3	138	1	ATPE_CYACA ATP SYNTHASE EPSILON C	1.14e+02
1	142	14.3	5	14.3	142	1	V16K_BLRV 16 KD PROTEIN (ORF2).	1.14e+02
1	149	14.3	5	14.3	149	1	RL2A_DROME 60S RIBOSOMAL PROTEIN	1.14e+02
1	165	14.3	5	14.3	165	1	TRY3_LUCU TRYP SIN ALPHA-3 (EC 3.	1.14e+02
1	180	14.3	5	14.3	180	1	YAF5_YEAST HYPOTHETICAL 19.9 KD P	1.14e+02
1	181	14.3	5	14.3	181	1	NQ06_THETH NADH-UBIQUINONE OXIDOR	1.14e+02
1	183	14.3	5	14.3	183	1	X391_METJA HYPOTHETICAL PROTEIN M	1.14e+02
1	185	14.3	5	14.3	185	1	YRK6_BACSU HYPOTHETICAL 20.7 KD P	1.14e+02
1	189	14.3	5	14.3	189	1	CHS1_ASPNG CHITIN SYNTHASE 1 (EC	1.14e+02
1	192	14.3	5	14.3	192	1	KAD_METJA ADENYLATE KINASE (EC 2	1.14e+02
1	197	14.3	5	14.3	197	1	SODF_PSEPU SUPEROXIDE DISMUTASE I	1.14e+02
1	197	14.3	5	14.3	197	1	HPAP_BURSO HRP ASSOCIATED PROTEIN	1.14e+02
1	205	14.3	5	14.3	205	1	YK07_YEAST HYPOTHETICAL 22.7 KD P	1.14e+02
1	208	14.3	5	14.3	208	1	INCB_ECOLI INCB PROTEIN.	1.14e+02
1	211	14.3	5	14.3	211	1	YOR3_SOUV3 HYPOTHETICAL 22.3 KD P	1.14e+02
1	212	14.3	5	14.3	212	1	COAT_PEBV COAT PROTEIN (CAPSID P	1.14e+02
1	218	14.3	5	14.3	218	1	RR3_TOBAC CHLOROPLAST 30S RIBOSO	1.14e+02
1	219	14.3	5	14.3	219	1	NUJC_SYNY3 NADH-UBIQUINONE OXIDOR	1.14e+02
1	219	14.3	5	14.3	219	1	TPIS_METJA PUTATIVE TRIOSEPHOSPHA	1.14e+02
1	220	14.3	5	14.3	220	1	NUOB_ECOLI NADH DEHYDROGENASE I C	1.14e+02
1	227	14.3	5	14.3	227	1	YOR2_EAV HYPOTHETICAL 25.6 KD P	1.14e+02
1	227	14.3	5	14.3	227	1	PIMT_BRARE PROTEIN-L-ISOASPARTATE	1.14e+02
1	228	14.3	5	14.3	228	1	CLCB_BOVIN CLATHRIN LIGHT CHAIN B	1.14e+02
1	228	14.3	5	14.3	228	1	Y516_BORBU HYPOTHETICAL TRNA/RNA	1.14e+02
1	228	14.3	5	14.3	228	1	CLCB_BOVIN CLATHRIN LIGHT CHAIN B	1.14e+02
1	229	14.3	5	14.3	229	1	SEM5_CAEEL SEX MUSCLE ABNORMAL PR	1.14e+02
1	230	14.3	5	14.3	230	1	VP26_NPVOP CLATHRIN LIGHT CHAIN B	1.14e+02
1	235	14.3	5	14.3	235	1	CLCA_MOUSE CLATHRIN LIGHT CHAIN A	1.14e+02
1	238	14.3	5	14.3	238	1	NEUM_HUMAN NEUROMODULIN (AXONAL M	1.14e+02
1	239	14.3	5	14.3	239	1	YGI1_YEAST HYPOTHETICAL 27.6 KD P	1.14e+02
1	239	14.3	5	14.3	239	1	DHAP_BOVIN ALDEHYDE DEHYDROGENASE	1.14e+02
1	239	14.3	5	14.3	239	1	UL20_HSVEB GENE 41 MEMBRANE PROTE	1.14e+02
1	243	14.3	5	14.3	243	1	YC22_METJA HYPOTHETICAL PROTEIN M	1.14e+02
1	243	14.3	5	14.3	243	1	CLCA_BOVIN CLATHRIN LIGHT CHAIN A	1.14e+02
1	245	14.3	5	14.3	245	1	CLCA_ANASP NADH-UBIQUINONE OXIDOR	1.14e+02
1	248	14.3	5	14.3	248	1	NUKC_SYNY3 NADH-UBIQUINONE OXIDOR	1.14e+02
1	248	14.3	5	14.3	248	1	CLCA_HUMAN CLATHRIN LIGHT CHAIN A	1.14e+02
1	250	14.3	5	14.3	250	1	GCH1_HUMAN GTP CYCLOHYDROLASE I (	1.14e+02
1	255	14.3	5	14.3	255	1	TRY4_LUCU TRYP SIN ALPHA-4 PRECUR	1.14e+02
1	255	14.3	5	14.3	255	1	PLSC_NEIGO 1-ACYL-SN-GLYCEROL-3-P	1.14e+02
1	263	14.3	5	14.3	263	1	YMF_PECOLI HYPOTHETICAL 28.2 KD P	1.14e+02
1	263	14.3	5	14.3	263	1	YXR_PORPU HYPOTHETICAL 28.3 KD P	1.14e+02
1	265	14.3	5	14.3	265	1	APAL_TUPGB APOLIPOPROTEIN A-I PRE	1.14e+02
1	265	14.3	5	14.3	265	1	HME2_BRARE HOMEOBOX PROTEIN ENGRA	1.14e+02
1	267	14.3	5	14.3	267	1	YCHA_HAEIN HYPOTHETICAL PROTEIN H	1.14e+02
1	270	14.3	5	14.3	270	1	YD1J_SCHPO HYPOTHETICAL 31.5 KD P	1.14e+02
1	271	14.3	5	14.3	271	1	IL1J_MACMU INTERLEUKIN-1 ALPHA PR	1.14e+02
1	272	14.3	5	14.3	272	1	CYNT_SYNP7 CARBONIC ANHYDRASE (EC	1.14e+02
1	275	14.3	5	14.3	275	1	NIF3_AZOVI NITROGENASE IRON PROTE	1.14e+02
1	279	14.3	5	14.3	279	1	Y43D_MYCPN HYPOTHETICAL LIPOPROTE	1.14e+02
1	281	14.3	5	14.3	281	1	Y054_METJA HYPOTHETICAL PROTEIN M	1.14e+02
1	281	14.3	5	14.3	281	1	BLAC_STRAU BETA-LACTAMASE PRECURS	1.14e+02
1	289	14.3	5	14.3	289	1	YADA_YEAST HYPOTHETICAL 33.2 KD P	1.14e+02
1	289	14.3	5	14.3	289	1	CRTB_THETH PHYTOENE SYNTHASE (EC	1.14e+02
1	294	14.3	5	14.3	294	1	APG5_YEAST AUTOPHAGY PROTEIN APG5	1.14e+02
1	296	14.3	5	14.3	296	1	RPIR_ECOLI RPIR PROTEIN.	1.14e+02
1	297	14.3	5	14.3	297	1	YGLB_YEAST HYPOTHETICAL 34.7 KD P	1.14e+02
1	300	14.3	5	14.3	300	1	ARGB_METJA PROBABLE ACETYLGLUTATH	1.14e+02
1	305	14.3	5	14.3	305	1	GSFC_VIBCH GENERAL SECRETION PATH	1.14e+02
1	308	14.3	5	14.3	308	1	YV53_HSVI1 HYPOTHETICAL GENE 53 P	1.14e+02
1	309	14.3	5	14.3	309	1	YC30_ODOSI HYPOTHETICAL TRANSCRIP	1.14e+02
1	311	14.3	5	14.3	311	1	CEFE_STRCL DEACETOXYCEPHALOSPORIN	1.14e+02
1	317	14.3	5	14.3	317	1	CEFE_NOCLA DEACETOXYCEPHALOSPORIN	1.14e+02
1	317	14.3	5	14.3	317	1	Y871_METJA HYPOTHETICAL PROTEIN M	1.14e+02
1	319	14.3	5	14.3	319	1	NBP1_YEAST NAP1-BINDING PROTEIN.	1.14e+02

97	5	14.3	321	1	GLK_BACSU	GLUCOKINASE (EC 2.7.1.	1.14e+02	170	1	HXA3_MOUSE	HOMEBOX PROTEIN HOX-A	1.14e+02
98	5	14.3	324	1	NODZ_BRAJA	MODULATION PROTEIN Z (	1.14e+02	171	1	NCAP_CVHOC	NUCLEOCAPSID PROTEIN.	1.14e+02
99	5	14.3	326	1	LXB2_PHOLE	ALKANAL MONOOXYGENASE	1.14e+02	172	1	ARGE_DICDI	ACETYLORNITHINE DEACET	1.14e+02
100	5	14.3	327	1	FASA_MOUSE	FASK RECEPTOR PRECURSO	1.14e+02	173	1	GSHR_BURCE	GLUTATHIONE REDUCTASE	1.14e+02
101	5	14.3	328	1	RPOA_HAEIN	DNA-DIRECTED RNA POLYM	1.14e+02	174	1	DHAP_RAT	ALDEHYDE DEHYDROGENASE	1.14e+02
102	5	14.3	328	1	NODZ_AZOCA	MODULATION PROTEIN Z (	1.14e+02	175	1	YOTA_CAEEL	HYPOTHETICAL 51.8 KD P	1.14e+02
103	5	14.3	329	1	YAZE_SCHPO	HYPOTHETICAL 36.7 KD P	1.14e+02	176	1	DHAP_MOUSE	ALDEHYDE DEHYDROGENASE	1.14e+02
104	5	14.3	331	1	ARGK_ECOLI	LAC/AO TRANSPORT SYSTE	1.14e+02	177	1	AAHP_ARATH	ASPARTATE AMINOTRANSFE	1.14e+02
105	5	14.3	332	1	EXPA_CEPAC	DEACETOXYCEPHALOSPORIN	1.14e+02	178	1	DHAP_HUMAN	ALDEHYDE DEHYDROGENASE	1.14e+02
106	5	14.3	333	1	BCHX_RHOC	CHLOROPHYLLIDE REDUCTA	1.14e+02	179	1	Y856_METTH	HYPOTHETICAL PROTEIN M	1.14e+02
107	5	14.3	333	1	BCHX_RHOSH	CHLOROPHYLLIDE REDUCTA	1.14e+02	180	1	Y041_METJA	HYPOTHETICAL PROTEIN M	1.14e+02
108	5	14.3	337	1	ASPO_PSES7	GLUTAMINASE-ASPARAGINA	1.14e+02	181	1	YSCP_YERPE	YOP PROTEINS TRANSLOCA	1.14e+02
109	5	14.3	339	1	EBAL_FLAME	ENDO-BETA-N-ACETYLGLUC	1.14e+02	182	1	PGKC_CRIFA	PHOSPHOGLYCERATE KINAS	1.14e+02
110	5	14.3	339	1	LICA_HAEIN	LICA PROTEIN.	1.14e+02	183	1	PUR8_ECOLI	ADENYLOSUCCINATE LYASE	1.14e+02
111	5	14.3	340	1	GBX2_CHICK	HOMEBOX PROTEIN GRX-2	1.14e+02	184	1	PUR8_HAEIN	ADENYLOSUCCINATE LYASE	1.14e+02
112	5	14.3	343	1	VANA_ENTFC	VANCOMYCIN/TEICOPLANIN	1.14e+02	185	1	VIPR_HUMAN	VASOACTIVE INTESTINAL	1.14e+02
113	5	14.3	343	1	NIDO_MOUSE	SODIUM-DEPENDENT DOPAM	1.14e+02	186	1	VIPR_PIG	VASOACTIVE INTESTINAL	1.14e+02
114	5	14.3	345	1	APOH_MOUSE	BETA-2-GLYCOPROTEIN I	1.14e+02	187	1	CG24_YEAST	G2/MITOTIC-SPECIFIC CY	1.14e+02
115	5	14.3	346	1	YN53_CABEL	HYPOTHETICAL 38.8 KD P	1.14e+02	188	1	VL2_HPV55	MINOR CAPSID PROTEIN L	1.14e+02
116	5	14.3	346	1	TAS_ECOLI	TAS PROTEIN.	1.14e+02	189	1	VE2_HPV09	REGULATORY PROTEIN E2.	1.14e+02
117	5	14.3	347	1	VALL_SLCV	ALL PROTEIN.	1.14e+02	190	1	YKX3_YEAST	HYPOTHETICAL 54.4 KD P	1.14e+02
118	5	14.3	355	1	YM79_YEAST	HYPOTHETICAL 37.4 KD P	1.14e+02	191	1	N2B_HAEIR	PUTATIVE ATPASE N2B (H	1.14e+02
119	5	14.3	360	1	DFR4_DIACA	DHYDROFLAVONOL-4-REDU	1.14e+02	192	1	EM55_MOUSE	55 KD ERYTHROCYTE MEMB	1.14e+02
120	5	14.3	361	1	YP28_METTF	HYPOTHETICAL 40.6 KD P	1.14e+02	193	1	EM55_HUMAN	55 KD ERYTHROCYTE MEMB	1.14e+02
121	5	14.3	370	1	NRFF_HAEIN	CYTOCHROME C-TYPE BIOG	1.14e+02	194	1	FIBB_BOVIN	FIBRINOGEN BETA CHAIN.	1.14e+02
122	5	14.3	370	1	ATPG_ODOSI	ATP SYNTHASE GAMMA CHA	1.14e+02	195	1	SP5R_BACSU	STAGE V SPOULATION PR	1.14e+02
123	5	14.3	372	1	NTE3_TOBAC	MITOGEN-ACTIVATED PROT	1.14e+02	196	1	DHAT_HUMAN	ALDEHYDE DEHYDROGENASE	1.14e+02
124	5	14.3	373	1	LUXE_PHOLE	LONG-CHAIN-FATTY-ACID-	1.14e+02	197	1	S61A_RAT	PROTEIN TRANSPORT PROT	1.14e+02
125	5	14.3	373	1	FLYD_ASPNG	PECTIN LASE D PRECURS	1.14e+02	198	1	S61A_CANFA	PROTEIN TRANSPORT PROT	1.14e+02
126	5	14.3	382	1	VEG0_YEAST	HYPOTHETICAL ZINC-TYPE	1.14e+02	199	1	RBL_PHYPA	RIBULOSE BISPHTOSPHATE	1.14e+02
127	5	14.3	383	1	YAG2_HPV03	REGULATORY PROTEIN E2.	1.14e+02	200	1	COGY_XENLA	STROMELYSIN-3 PRECURSO	1.14e+02
128	5	14.3	384	1	MAPK_PETHY	MITOGEN-ACTIVATED PROT	1.14e+02	201	1	PGKC_LEIMA	PHOSPHOGLYCERATE KINAS	1.14e+02
129	5	14.3	385	1	DH48_HUMAN	ALDEHYDE DEHYDROGENASE	1.14e+02	202	1	PGKC_LEIME	PHOSPHOGLYCERATE KINAS	1.14e+02
130	5	14.3	386	1	ADT2_SOLITU	ADP,ATP CARRIER PROTEI	1.14e+02	203	1	VGIC_PRVIF	GLYCOPROTEIN GIII PREC	1.14e+02
131	5	14.3	387	1	YGH5_YEAST	HYPOTHETICAL 44.6 KD P	1.14e+02	204	1	FASC_XENLA	FASCIN.	1.14e+02
132	5	14.3	389	1	YPT6_CAEEL	HYPOTHETICAL 43.5 KD P	1.14e+02	205	1	DHAL_PSEOL	ALDEHYDE DEHYDROGENASE	1.14e+02
133	5	14.3	390	1	P2CB_RAT	PROTEIN PHOSPHATASE 2C	1.14e+02	206	1	DH4A_MOUSE	FATTY ALDEHYDE DEHYDRO	1.14e+02
134	5	14.3	390	1	ADHB_BACSU	PROBABLE NADH-DEPENDEN	1.14e+02	207	1	DH4A_RAT	FATTY ALDEHYDE DEHYDRO	1.14e+02
135	5	14.3	390	1	P2CB_MOUSE	PROTEIN PHOSPHATASE 2C	1.14e+02	208	1	DH4A_HUMAN	FATTY ALDEHYDE DEHYDRO	1.14e+02
136	5	14.3	391	1	DP3B_SYNY3	DNA POLYMERASE IIII, BE	1.14e+02	209	1	GUN1_BACSA	ENDOGUCANASE A (EC 3.	1.14e+02
137	5	14.3	392	1	YKH7_YEAST	HYPOTHETICAL 46.0 KD P	1.14e+02	210	1	HMS1_YEAST	ZINC FINGER PROTEIN HM	1.14e+02
138	5	14.3	392	1	TEBB_STVIMY	TELOMERE-BINDING PROTE	1.14e+02	211	1	GLPK_BACSU	GLYCEROL KINASE (EC 2.	1.14e+02
139	5	14.3	399	1	NAGA_VIBFU	N-ACETYLGLUCOSAMINE-6-	1.14e+02	212	1	BAF1_KLUMA	TRANSCRIPTION FACTOR B	1.14e+02
140	5	14.3	402	1	VGLD_PVRVI	GLYCOPROTEIN GP50.	1.14e+02	213	1	RP54_PSEAE	RNA POLYMERASE SIGMA-5	1.14e+02
141	5	14.3	403	1	THIM_CANTR	ACETYL-COA ACETVLTANS	1.14e+02	214	1	CBS_DICDI	CYSTATHIONINE BETA-SYN	1.14e+02
142	5	14.3	406	1	P2C_LEICH	PROTEIN PHOSPHATASE 2C	1.14e+02	215	1	CISA_BACSU	PUTATIVE DNA RECOMBINA	1.14e+02
143	5	14.3	408	1	NISH_MYCPN	NIFS-LIKE PROTEIN.	1.14e+02	216	1	YM05_YEAST	HYPOTHETICAL 58.0 KD P	1.14e+02
144	5	14.3	409	1	R23B_HUMAN	UV EXCISION REPAIR PRO	1.14e+02	217	1	AFX1_HUMAN	PUTATIVE FORK HEAD DOM	1.14e+02
145	5	14.3	409	1	GUN2_BACSA	ENDOGUCANASE B (EC 3.	1.14e+02	218	1	MTBR_BPRH1	MODIFICATION METHYLASE	1.14e+02
146	5	14.3	409	1	ARR2_HUMAN	BETA-ARRESTIN 2.	1.14e+02	219	1	PGKA_TRYBB	PHOSPHOGLYCERATE KINAS	1.14e+02
147	5	14.3	410	1	ARR2_RAT	BETA-ARRESTIN 2.	1.14e+02	220	1	PGKA_CRIFA	PHOSPHOGLYCERATE KINAS	1.14e+02
148	5	14.3	412	1	YLS7_YEAST	PUTATIVE DIOXYGENASE Y	1.14e+02	221	1	ACHG_TORCA	ACETYLCHOLINE RECEPTOR	1.14e+02
149	5	14.3	416	1	ELTI_CAEEL	TRANSCRIPTION FACTOR E	1.14e+02	222	1	PGKD_TRYBB	PHOSPHOGLYCERATE KINAS	1.14e+02
150	5	14.3	416	1	FILA_HUMAN	FILAGRIN PRECURSOR (F	1.14e+02	223	1	PGKG_TRYCO	PHOSPHOGLYCERATE KINAS	1.14e+02
151	5	14.3	417	1	PGKB_CRIFA	PHOSPHOGLYCERATE KINAS	1.14e+02	224	1	ATPA_BUCAP	ATP SYNTHASE ALPHA CHA	1.14e+02
152	5	14.3	417	1	PGKB_LEIMA	PHOSPHOGLYCERATE KINAS	1.14e+02	225	1	VGIG_VSVSJ	SPIKE GLYCOPROTEIN PRE	1.14e+02
153	5	14.3	417	1	PGKB_LEIME	PHOSPHOGLYCERATE KINAS	1.14e+02	226	1	VGIG_VSVIG	SPIKE GLYCOPROTEIN PRE	1.14e+02
154	5	14.3	417	1	LII0_CABEL	LIN-10 PROTEIN.	1.14e+02	227	1	Y079_MYCTO	HYPOTHETICAL 53.9 KD P	1.14e+02
155	5	14.3	419	1	SECY_PAVLU	PREPROTEIN TRANSLOCASE	1.14e+02	228	1	VGIG_HSVBC	GLYCOPROTEIN GIII PREC	1.14e+02
156	5	14.3	419	1	SEP2_DROME	SEPTIN 2.	1.14e+02	229	1	MET3_YEAST	SULFATE ADENYLYLTRANSF	1.14e+02
157	5	14.3	420	1	ARR2_BOVIN	BETA-ARRESTIN 2 (ARRS	1.14e+02	230	1	CLOS_CLOHI	ALPHA-CLOSTRIPAIN PREC	1.14e+02
158	5	14.3	420	1	PGKE_TRYBB	PHOSPHOGLYCERATE KINAS	1.14e+02	231	1	LEU1_ANASP	2-ISOPROPYLMALATE SYNT	1.14e+02
159	5	14.3	420	1	PGK1_TRYCO	PHOSPHOGLYCERATE KINAS	1.14e+02	232	1	EGR1_MOUSE	EARLY GROWTH RESPONSE	1.14e+02
160	5	14.3	421	1	ZP3_BOVIN	ZONA PELLUCIDA SPERM-B	1.14e+02	233	1	CN4C_RAT	CAMP-DEPENDENT 3',5'-C	1.14e+02
161	5	14.3	421	1	PGKB_TRYBB	PHOSPHOGLYCERATE KINAS	1.14e+02	234	1	PNUT_DROME	PEANUT PROTEIN.	1.14e+02
162	5	14.3	424	1	MTB1_BACSH	MODIFICATION METHYLASE	1.14e+02	235	1	PPB_LYSEN	ALKALINE PHOSPHATASE P	1.14e+02
163	5	14.3	427	1	YBHC_ECOLI	HYPOTHETICAL 46.1 KD P	1.14e+02	236	1	GUX1_ASPAC	EXOGUCANASE I PRECURS	1.14e+02
164	5	14.3	427	1	VG24_BPT4	HEAD VERTEX PROTEIN GP	1.14e+02	237	1	MIG1_KLUMA	REGULATORY PROTEIN MIG	1.14e+02
165	5	14.3	430	1	PBX2_MOUSE	PRE-B-CELL LEUKEMIA TR	1.14e+02	238	1	ODP2_ACHLA	DIHYDROLIPOAMIDE ACETY	1.14e+02
166	5	14.3	430	1	PBX3_HUMAN	PRE-B-CELL LEUKEMIA TR	1.14e+02	239	1	HBPA_HAEIN	HEME-BINDING PROTEIN A	1.14e+02
167	5	14.3	432	1	RAD9_SCHOT	DNA REPAIR PROTEIN RAD	1.14e+02	240	1	LYE2_RAT	E-SELECTIN PRECURSOR (	1.14e+02
168	5	14.3	435	1	FMRA_ATEL	ANTHO-REFAMIDE NEUROPEP	1.14e+02	241	1	YFE2_YEAST	HYPOTHETICAL 62.5 KD P	1.14e+02
169	5	14.3	440	1	PGKC_TRYBB	PHOSPHOGLYCERATE KINAS	1.14e+02	242	1	EG10_CAEEL	REGULATOR OF G-PROTEIN	1.14e+02



HYPOTHEICAL 107.7 KD	1.14e+02
GLUTAMATE-AMMONIA-LIGA	1.14e+02
GLUTAMATE RECEPTOR, IO	1.14e+02
HYPOTHEICAL 104.1 KD	1.14e+02
EUKARYOTIC INITIATION	1.14e+02
PRUVATE, PHOSPHATE DIK	1.14e+02
PRUVATE, PHOSPHATE DIK	1.14e+02
HYPOTHEICAL 108.7 KD	1.14e+02
GENERAL VESICULAR TRAN	1.14e+02
PROTIVILIN (100 KD AC	1.14e+02
GENERAL VESICULAR TRAN	1.14e+02
TRANSCRIPTIONAL ACTIVA	1.14e+02
C-1-TETRAHYDROFOLATE S	1.14e+02
P115 PROTEIN HOMOLOG.	1.14e+02
CALCIUM-TRANSPORTING A	1.14e+02
LYSOSOMAL ALPHA-MANNOS	1.14e+02
CALCIUM-TRANSPORTING A	1.14e+02
HYPOTHEICAL 113.4 KD	1.14e+02
TRANSLATION INITIATION	1.14e+02
SN12 PROTEIN (SR077 PR	1.14e+02
ALPHA, ALPHA-TREHALOSE-	1.14e+02
COMPLEMENT RECEPTOR	1.14e+02
P53-BINDING PROTEIN 53	1.14e+02
NEURONAL KINESIN HEAVY	1.14e+02
COMPLEMENT RECEPTOR	1.14e+02
SN11 PROTEIN (SR07 PRO	1.14e+02
CATION-TRANSPORTING AT	1.14e+02
BO12 PROTEIN (BEB1 PRO	1.14e+02
HYPOTHEICAL 117.3 KD	1.14e+02
USN1 PROTEIN.	1.14e+02
PHOSPHATIDYLINOSITOL 4	1.14e+02
PROTEIN KINASE C-LIKE	1.14e+02
PROTEIN KINASE C-LIKE	1.14e+02
FURIN-LIKE PROTEASE 1,	1.14e+02
DELTA-CRYSTALLIN ENHAN	1.14e+02
HYPOTHEICAL 127.3 KD	1.14e+02
REPETIN.	1.14e+02
HYPOTHEICAL 128.8 KD	1.14e+02
PROTEIN KINASE C-LIKE	1.14e+02
ASK10 PROTEIN.	1.14e+02
POTASSIUM CHANNEL PROT	1.14e+02
1-PHOSPHATIDYLINOSITOL	1.14e+02
NUCLEAR ENVELOPE PORE	1.14e+02
1-PHOSPHATIDYLINOSITOL	1.14e+02
1-PHOSPHATIDYLINOSITOL	1.14e+02
HYPOTHEICAL 136.3 KD	1.14e+02
DNA-DIRECTED RNA POLYM	1.14e+02
HYPOTHEICAL 143.6 KD	1.14e+02
DNA POLYMERASE (EC 2.7	1.14e+02
HYPOTHEICAL 133.5 KD	1.14e+02
HYPOTHEICAL PROTEIN M	1.14e+02
DNA-DIRECTED RNA POLYM	1.14e+02
TRANS-ACCTING TRANSCRIP	1.14e+02
BETA-N-ACETYLUHOSAMIN	1.14e+02
DNA-DIRECTED RNA POLYM	1.14e+02
HYPOTHEICAL 150.8 KD	1.14e+02
PROBABLE MAJOR GLYCOPR	1.14e+02
XANTHINE DEHYDROGENASE	1.14e+02
DNA-DIRECTED RNA POLYM	1.14e+02
HYPOTHEICAL 157.0 KD	1.14e+02
240 KD PROTEIN OF ROD	1.14e+02
EUKARYOTIC TRANSLATION	1.14e+02
SERRATE PROTEIN PRECUR	1.14e+02
REGULATORY PROTEIN CAT	1.14e+02
HYPOTHEICAL 161.2 KD	1.14e+02
AMPHOLIPULANASE PRECU	1.14e+02
AF-6 PROTEIN.	1.14e+02
VITELLOGENIN 6 PRECURS	1.14e+02
INTEGRIN ALPHA CHAIN-L	1.14e+02
VITELLOGENIN I PRECURS	1.14e+02
TENASCIN PRECURSOR (TN	1.14e+02
ADENYLATE CYCLASE (EC	1.14e+02
HYPOTHEICAL PROTEIN M	1.14e+02

389	5	14.3	1942	1	Y054_HUMAN	HYPOTHETICAL PROTEIN K	1.14e+02	462	160	1	YS57_CAEEL	HYPOTHETICAL CALCIUM-B	2.67e+03
390	5	14.3	2014	1	XJ07_YEAST	HYPOTHETICAL 22.9 KD	1.14e+02	463	160	1	YEM0_YEAST	HYPOTHETICAL 18.3 KD P	2.67e+03
391	5	14.3	2145	1	CYAA_PODAN	ADENYLATE CYCLASE (EC	1.14e+02	464	164	1	Y646_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
392	5	14.3	2164	1	POLG_HRY89	GENOME POLYPROTEIN [CO	1.14e+02	465	167	1	Y6G3_YEAST	HYPOTHETICAL 19.0 KD P	2.67e+03
393	5	14.3	2279	1	COAC_SCHPO	ACTYL-COA CARBOXYLASE	1.14e+02	466	168	1	Y168_ADE02	HYPOTHETICAL PROTEIN C	2.67e+03
394	5	14.3	2300	1	CYAA_NEUCR	ADENYLATE CYCLASE (EC	1.14e+02	467	169	1	Y769_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
395	5	14.3	2483	1	PCX_DROME	PENANEX PROTEIN.	1.14e+02	468	169	1	YKHO_YEAST	HYPOTHETICAL 19.8 KD P	2.67e+03
396	5	14.3	2491	1	TALA_DICDI	FILOPODIN (TALIN HOMOL	1.14e+02	469	169	1	X193_ECOLI	X POLYPEPTIDE (ORF 19)	2.67e+03
397	5	14.3	2843	1	APC_HUMAN	ADENOMATOUS POLYPOSIS	1.14e+02	470	169	1	X192_ECOLI	X POLYPEPTIDE (ORF 19)	2.67e+03
398	5	14.3	3033	1	POLG_HCVJ6	GENOME POLYPROTEIN [CO	1.14e+02	471	170	1	Y019_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03
399	5	14.3	3866	1	HRX_MOUSE	ZINC FINGER PROTEIN HR	1.14e+02	472	175	1	YPG_BWTVF	PUTATIVE GENOME-LINKED	2.67e+03
400	5	14.3	3924	1	ANKB_HUMAN	ANKYRIN, BRAIN VARIANT	1.14e+02	473	175	1	YPG_BWTVF	PUTATIVE GENOME-LINKED	2.67e+03
401	5	14.3	4393	1	PGBM_HUMAN	BASEMENT MEMBRANE-SPEC	1.14e+02	474	178	1	Y449_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
402	4	11.4	445	1	VNSX_CVBM	NONSTRUCTURAL 4.8 KD P	2.67e+03	475	179	1	Y281_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
403	4	11.4	57	1	YOR3_TTV1	HYPOTHETICAL 6.9 KD PR	2.67e+03	476	180	1	XG_HUMAN	XG GLYCOPROTEIN PRECUR	2.67e+03
404	4	11.4	60	1	YH05_VACV	HYPOTHETICAL 5.7 KD HI	2.67e+03	477	181	1	Y44S_RHISN	HYPOTHETICAL 20.0 KD P	2.67e+03
405	4	11.4	60	1	YL13_ADE41	HYPOTHETICAL 6.9 KD PR	2.67e+03	478	181	1	YH01_YEAST	HYPOTHETICAL 21.0 KD P	2.67e+03
406	4	11.4	61	1	VP12_ASEB7	VIRUS ATTACHMENT PROTE	2.67e+03	479	185	1	YCX8_BACSU	HYPOTHETICAL 22.0 KD P	2.67e+03
407	4	11.4	71	1	Y16K_BPT4	HYPOTHETICAL 8.1 KD PR	2.67e+03	480	187	1	YQGN_BACSU	HYPOTHETICAL 21.4 KD P	2.67e+03
408	4	11.4	71	1	YFX1_RHLE	HYPOTHETICAL 7.9 KD PR	2.67e+03	481	187	1	YDAL_ECOLI	HYPOTHETICAL 21.5 KD P	2.67e+03
409	4	11.4	74	1	WDNM_RAT	WDNM1 PROTEIN PRECURSO	2.67e+03	482	194	1	YOV8_CAEEL	HYPOTHETICAL 23.0 KD P	2.67e+03
410	4	11.4	75	1	YCX1_PORPU	HYPOTHETICAL 8.3 KD PR	2.67e+03	483	194	1	YGG8_YEAST	PUTATIVE MITOCHONDRIAL	2.67e+03
411	4	11.4	76	1	Y49L_SYNF3	YCF49-LIKE PROTEIN.	2.67e+03	484	196	1	ZG64_XENLA	GASTRULA ZINC FINGER P	2.67e+03
412	4	11.4	81	1	Y78A_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	485	196	1	ZG3_XENLA	GASTRULA ZINC FINGER P	2.67e+03
413	4	11.4	84	1	YE96_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	486	197	1	Y916_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
414	4	11.4	87	1	Y07A_BPT4	HYPOTHETICAL 10.2 KD P	2.67e+03	487	198	1	XA1_XENLA	XA-1 PROTEIN PRECURSOR	2.67e+03
415	4	11.4	90	1	Y105_BPT3	HYPOTHETICAL GENE 1.05	2.67e+03	488	201	1	YMBD_YEAST	HYPOTHETICAL 23.0 KD P	2.67e+03
416	4	11.4	91	1	YJFY_ECOLI	HYPOTHETICAL 10.1 KD P	2.67e+03	489	201	1	YIAS_BACSU	HYPOTHETICAL 22.0 KD P	2.67e+03
417	4	11.4	95	1	YS05_MOUSE	HYPOTHETICAL PROTEIN L	2.67e+03	490	202	1	YR44_CAEEL	HYPOTHETICAL 23.8 KD P	2.67e+03
418	4	11.4	96	1	YRS5_CAEEL	HYPOTHETICAL 11.1 KD P	2.67e+03	491	203	1	YPT1_NEUCR	GTP-BINDING PROTEIN YP	2.67e+03
419	4	11.4	103	1	YD29_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	492	204	1	YQF7_CAEEL	HYPOTHETICAL 22.4 KD P	2.67e+03
420	4	11.4	106	1	Y115_ADE07	HYPOTHETICAL 11.5 KD E	2.67e+03	493	205	1	Y119_YEAST	HYPOTHETICAL 23.9 KD P	2.67e+03
421	4	11.4	106	1	Y116_ADE02	HYPOTHETICAL 11.6 KD E	2.67e+03	494	206	1	Y21K_MSVN	HYPOTHETICAL 21.7 KD P	2.67e+03
422	4	11.4	109	1	Y353_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	495	210	1	Y753_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
423	4	11.4	109	1	VNS2_CVBM	NONSTRUCTURAL PROTEIN	2.67e+03	496	211	1	YNQ8_CAEEL	HYPOTHETICAL 23.7 KD P	2.67e+03
424	4	11.4	110	1	Y1S2_STRCO	INSERTION ELEMENT IS11	2.67e+03	497	212	1	Y331_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
425	4	11.4	112	1	WHEP_CAMDR	WHEY PROTEIN.	2.67e+03	498	212	1	YJFM_ECOLI	HYPOTHETICAL 23.4 KD P	2.67e+03
426	4	11.4	112	1	YCX6_CHLRE	HYPOTHETICAL 12.3 KD P	2.67e+03	499	213	1	YIN2_YEAST	HYPOTHETICAL 25.0 KD P	2.67e+03
427	4	11.4	112	1	Y143_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	500	214	1	YNV3_YEAST	HYPOTHETICAL 25.3 KD P	2.67e+03
428	4	11.4	113	1	YKGI_YEAST	HYPOTHETICAL 13.5 KD P	2.67e+03	501	214	1	YDL8_SCHPO	HYPOTHETICAL 23.9 KD P	2.67e+03
429	4	11.4	116	1	Y137_METH	HYPOTHETICAL PROTEIN M	2.67e+03	502	215	1	YKL7_YEAST	24.1 KD PROTEIN IN VMA	2.67e+03
430	4	11.4	117	1	YCC1_YEAST	HYPOTHETICAL 13.6 KD P	2.67e+03	503	216	1	YWR5_CAEEL	HYPOTHETICAL 25.1 KD P	2.67e+03
431	4	11.4	122	1	YQ09_YEAST	HYPOTHETICAL 14.0 KD P	2.67e+03	504	218	1	Y381_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
432	4	11.4	123	1	YPPE_BACSU	HYPOTHETICAL 14.5 KD P	2.67e+03	505	218	1	YB9J_YEAST	HYPOTHETICAL RAS-RELAT	2.67e+03
433	4	11.4	123	1	Y055_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	506	221	1	YFH8_YEAST	HYPOTHETICAL 25.9 KD P	2.67e+03
434	4	11.4	124	1	YG78_BRP21	HYPOTHETICAL 13.5 KD P	2.67e+03	507	222	1	Y101_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03
435	4	11.4	124	1	Y729_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	508	222	1	YKGH_ECOLI	HYPOTHETICAL 25.6 KD P	2.67e+03
436	4	11.4	129	1	YNXA_CAEEL	HYPOTHETICAL 14.7 KD P	2.67e+03	509	224	1	YF08_METJA	HYPOTHETICAL ABC TRANS	2.67e+03
437	4	11.4	132	1	YBL2_STRCI	HYPOTHETICAL 14.2 KD P	2.67e+03	510	225	1	Y059_SYNF3	HYPOTHETICAL 24.9 KD P	2.67e+03
438	4	11.4	132	1	YAF0_ECOLI	HYPOTHETICAL 15.5 KD P	2.67e+03	511	226	1	Y069_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
439	4	11.4	133	1	YHGG_ECOLI	HYPOTHETICAL 15.1 KD P	2.67e+03	512	226	1	Y703_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
440	4	11.4	136	1	YKJ7_YEAST	HYPOTHETICAL 15.4 KD P	2.67e+03	513	228	1	Y132_NPVOP	HYPOTHETICAL 26.0 KD P	2.67e+03
441	4	11.4	139	1	Y28_BPT7	HYPOTHETICAL GENE 2.8	2.67e+03	514	230	1	YGOJ_ECOLI	HYPOTHETICAL 25.5 KD P	2.67e+03
442	4	11.4	140	1	YP15_STRAU	HYPOTHETICAL 15.5 KD P	2.67e+03	515	231	1	YSAL_YEAST	YSAL PROTEIN.	2.67e+03
443	4	11.4	140	1	Y337_MYCPN	HYPOTHETICAL 15.5 KD P	2.67e+03	516	232	1	Y529_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
444	4	11.4	141	1	YFE3_CLOPA	HYPOTHETICAL PROTEIN I	2.67e+03	517	232	1	Y4IK_RHISN	HYPOTHETICAL 26.8 KD P	2.67e+03
445	4	11.4	141	1	VTUL_DROME	VITELLINE MEMBRANE PRO	2.67e+03	518	235	1	YQ57_CAEEL	HYPOTHETICAL 27.9 KD P	2.67e+03
446	4	11.4	141	1	Y496_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	519	237	1	YMI7_YEAST	PROBABLE TRANSCRIPTION	2.67e+03
447	4	11.4	144	1	YD31_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	520	239	1	YC27_CVAPA	HYPOTHETICAL PROTEIN M	2.67e+03
448	4	11.4	146	1	YAAR_BACSU	HYPOTHETICAL 16.7 KD P	2.67e+03	521	241	1	YF45_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
449	4	11.4	146	1	YAT7_SCHPO	VERY HYPOTHETICAL 17.1	2.67e+03	522	241	1	Y293_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03
450	4	11.4	147	1	YDEF_SCHPO	HYPOTHETICAL 16.9 KD P	2.67e+03	523	242	1	YBGP_ECOLI	HYPOTHETICAL FMBRIAL	2.67e+03
451	4	11.4	148	1	Y452_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	524	243	1	Y004_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
452	4	11.4	150	1	YM8W_YEAST	HYPOTHETICAL 17.2 KD P	2.67e+03	525	243	1	YCGE_ECOLI	HYPOTHETICAL TRANSCRIP	2.67e+03
453	4	11.4	151	1	YPR3_ECOLI	HYPOTHETICAL 17.4 KD P	2.67e+03	526	244	1	YENR_YEREN	YENR REGULATORY PROTEI	2.67e+03
454	4	11.4	151	1	YROS_RHET	HYPOTHETICAL 16.8 KD P	2.67e+03	527	244	1	Y384_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
455	4	11.4	151	1	YA8B_SCHPO	HYPOTHETICAL 18.0 KD P	2.67e+03	528	245	1	Y1T8_YEAST	HYPOTHETICAL 28.4 KD P	2.67e+03
456	4	11.4	152	1	YK71_YEAST	HYPOTHETICAL 17.0 KD P	2.67e+03	529	246	1	YBEC_ECOLI	HYPOTHETICAL 26.4 KD P	2.67e+03
457	4	11.4	153	1	YPG_BYDVP	PUTATIVE GENOME-LINKED	2.67e+03	530	247	1	YH01_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
458	4	11.4	153	1	YN21_YEAST	HYPOTHETICAL 17.2 KD P	2.67e+03	531	248	1	Y4EK_RHISN	PUTATIVE SHORT-CHAIN T	2.67e+03
459	4	11.4	154	1	YDTE_SCHPO	HYPOTHETICAL 17.4 KD P	2.67e+03	532	248	1	YIFA_YEAST	YIPI PROTEIN.	2.67e+03
460	4	11.4	155	1	YK54_YEAST	HYPOTHETICAL 18.4 KD P	2.67e+03	533	249	1	YF82_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
461	4	11.4	157	1	WH16_STRCO	16.7 KD PROTEIN IN WHI	2.67e+03	534	251	1	Y44D_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03

535	4	11.4	252	1	YCAL_PLAFA	HYPOTHETICAL PROTEIN I	2.67e+03	608	4	11.4	322	1	YK02_YEAST	HYPOTHETICAL 36.6 KD P	2.67e+03
536	4	11.4	255	1	YPMR_BACSU	HYPOTHETICAL 28.9 KD P	2.67e+03	609	4	11.4	324	1	Y371_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
537	4	11.4	256	1	YC27_CYACA	PROBABLE TRANSCRIPTION	2.67e+03	610	4	11.4	324	1	Y700_METJA	HYPOTHETICAL 36.5 KD P	2.67e+03
538	4	11.4	256	1	Y448_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	611	4	11.4	325	1	Y371_MYCGE	HYPOTHETICAL 36.5 KD P	2.67e+03
539	4	11.4	258	1	YK20_METJA	HYPOTHETICAL 27.3 KD P	2.67e+03	612	4	11.4	326	1	Y299_CAEEL	HYPOTHETICAL 37.4 KD P	2.67e+03
540	4	11.4	261	1	YF02_ECOLI	HYPOTHETICAL 29.2 KD P	2.67e+03	613	4	11.4	328	1	YJ77_YEAST	HYPOTHETICAL LIPASE IN	2.67e+03
541	4	11.4	261	1	YF31_FRG3V	EARLY 31 KD PROTEIN	2.67e+03	614	4	11.4	329	1	YQ01_CAEEL	HYPOTHETICAL 39.0 KD P	2.67e+03
542	4	11.4	262	1	YXBG_BACSU	HYPOTHETICAL OXIDOREDU	2.67e+03	615	4	11.4	330	1	Y0A4_MYCTU	HYPOTHETICAL ABC TRANS	2.67e+03
543	4	11.4	262	1	YD05_CAEEL	HYPOTHETICAL 29.7 KD P	2.67e+03	616	4	11.4	331	1	YPLP_BACSU	PUTATIVE SIGMA L-DEPEN	2.67e+03
544	4	11.4	263	1	YBGF_ECOLI	HYPOTHETICAL 28.5 KD P	2.67e+03	617	4	11.4	333	1	Y237_SYNY3	HYPOTHETICAL SUGAR KIN	2.67e+03
545	4	11.4	263	1	Y07K_MYCTU	HYPOTHETICAL 28.5 KD P	2.67e+03	618	4	11.4	334	1	Y009_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03
546	4	11.4	267	1	YTXD_BACME	HYPOTHETICAL 29.3 KD P	2.67e+03	619	4	11.4	334	1	YH05_YEAST	HYPOTHETICAL 37.9 KD P	2.67e+03
547	4	11.4	267	1	YVST_SFVS	NONSTRUCTURAL PROTEIN	2.67e+03	620	4	11.4	338	1	Y941_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
548	4	11.4	267	1	YAFB_ECOLI	HYPOTHETICAL OXIDOREDU	2.67e+03	621	4	11.4	340	1	YAKC_SCHPO	HYPOTHETICAL 37.7 KD P	2.67e+03
549	4	11.4	269	1	YD08_SCHPO	HYPOTHETICAL 30.5 KD P	2.67e+03	622	4	11.4	343	1	YMD4_CAEEL	HYPOTHETICAL 40.8 KD P	2.67e+03
550	4	11.4	269	1	YR05_CAEEL	HYPOTHETICAL 30.7 KD P	2.67e+03	623	4	11.4	343	1	YQKA_BACSU	HYPOTHETICAL 39.0 KD P	2.67e+03
551	4	11.4	270	1	YKAA_CAEEL	HYPOTHETICAL 31.2 KD P	2.67e+03	624	4	11.4	346	1	YC89_ARCFU	HYPOTHETICAL PROTEIN A	2.67e+03
552	4	11.4	270	1	YLEU_CORGL	HYPOTHETICAL 29.6 KD P	2.67e+03	625	4	11.4	346	1	Y343_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
553	4	11.4	270	1	Y442_MYCGE	HYPOTHETICAL PROTEIN P	2.67e+03	626	4	11.4	347	1	Y241_CAEEL	HYPOTHETICAL 39.0 KD P	2.67e+03
554	4	11.4	271	1	YDJO_ECOLI	HYPOTHETICAL 30.7 KD P	2.67e+03	627	4	11.4	348	1	Y351_STREF	HYPOTHETICAL 35.5 KD P	2.67e+03
555	4	11.4	273	1	YF30_METTM	HYPOTHETICAL 30.7 KD P	2.67e+03	628	4	11.4	348	1	YLM1_CAEEL	HYPOTHETICAL 41.0 KD P	2.67e+03
556	4	11.4	274	1	Y179_MYCGE	HYPOTHETICAL ABC TRANS	2.67e+03	629	4	11.4	348	1	YKDT_BACSU	PHAGE-LIKE ELEMENT PBS	2.67e+03
557	4	11.4	275	1	YAAAT_BACSU	HYPOTHETICAL 31.2 KD P	2.67e+03	630	4	11.4	351	1	YNO4_YEAST	HYPOTHETICAL 40.0 KD P	2.67e+03
558	4	11.4	276	1	Y44C_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03	631	4	11.4	352	1	YORM_TTV1	HYPOTHETICAL 38.6 KD P	2.67e+03
559	4	11.4	276	1	Y842_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	632	4	11.4	354	1	YXSA_ANASP	EXCISEASE A (XISA) PROTE	2.67e+03
560	4	11.4	277	1	YAC3_YEAST	HYPOTHETICAL 32.3 KD P	2.67e+03	633	4	11.4	355	1	YX44_YEAST	HYPOTHETICAL 40.7 KD G	2.67e+03
561	4	11.4	278	1	YQHA_BACSU	HYPOTHETICAL 31.8 KD P	2.67e+03	634	4	11.4	355	1	Y121_TREPA	HYPOTHETICAL PROTEIN T	2.67e+03
562	4	11.4	278	1	Y751_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	635	4	11.4	358	1	Y074_METJA	HYPOTHETICAL ATP-BINDI	2.67e+03
563	4	11.4	279	1	Y511_CAEEL	HYPOTHETICAL 30.7 KD P	2.67e+03	636	4	11.4	358	1	YC07_KLEPN	HYPOTHETICAL 41.2 KD P	2.67e+03
564	4	11.4	280	1	Y246_XENLA	GASTRULA ZINC FINGER P	2.67e+03	637	4	11.4	359	1	Y4UE_RHISN	PUTATIVE TRANSPOSASE Y	2.67e+03
565	4	11.4	281	1	YP41_ROTNR	OUTER CAPSID PROTEIN V	2.67e+03	638	4	11.4	359	1	Y862_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
566	4	11.4	282	1	YH01_YEAST	HYPOTHETICAL 32.1 KD P	2.67e+03	639	4	11.4	359	1	YH09_YEAST	HYPOTHETICAL 39.4 KD P	2.67e+03
567	4	11.4	283	1	Y43C_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03	640	4	11.4	364	1	YH05_YEAST	HYPOTHETICAL 40.9 KD P	2.67e+03
568	4	11.4	285	1	Y007_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03	641	4	11.4	364	1	WNT6_MOUSE	WNT-6 PROTEIN PRECURSO	2.67e+03
569	4	11.4	285	1	YDUA_ECOLI	HYPOTHETICAL 32.5 KD P	2.67e+03	642	4	11.4	368	1	Y974_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
570	4	11.4	285	1	YHMF_METFE	HYPOTHETICAL 32.2 KD P	2.67e+03	643	4	11.4	369	1	Y519_CAEEL	HYPOTHETICAL 42.1 KD P	2.67e+03
571	4	11.4	286	1	YBL2_YEAST	HYPOTHETICAL 32.7 KD P	2.67e+03	644	4	11.4	369	1	YAD6_YEAST	HYPOTHETICAL 43.2 KD P	2.67e+03
572	4	11.4	289	1	Y973_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	645	4	11.4	372	1	YEF1_ECOLI	HYPOTHETICAL 43.2 KD P	2.67e+03
573	4	11.4	289	1	YXAB_BACSU	HYPOTHETICAL 31.8 KD P	2.67e+03	646	4	11.4	373	1	Y812_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
574	4	11.4	290	1	Y43A_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03	647	4	11.4	374	1	YMP4_CAEEL	HYPOTHETICAL 42.9 KD P	2.67e+03
575	4	11.4	290	1	Y070_NPVAC	HYPOTHETICAL 34.4 KD P	2.67e+03	648	4	11.4	374	1	YG32_AQUAE	HYPOTHETICAL PROTEIN A	2.67e+03
576	4	11.4	291	1	YC38_PORPU	HYPOTHETICAL 32.4 KD P	2.67e+03	649	4	11.4	374	1	Y158_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
577	4	11.4	292	1	YOPN_YERPS	OUTER MEMBRANE PROTEIN	2.67e+03	650	4	11.4	375	1	Y021_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
578	4	11.4	293	1	YOBQ_MYCTU	HYPOTHETICAL 30.8 KD P	2.67e+03	651	4	11.4	375	1	YH01_ECOLI	HYPOTHETICAL 43.1 KD P	2.67e+03
579	4	11.4	293	1	Y326_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03	652	4	11.4	378	1	YK84_CAEEL	HYPOTHETICAL 43.0 KD P	2.67e+03
580	4	11.4	293	1	YNP7_CAEEL	HYPOTHETICAL 33.7 KD P	2.67e+03	653	4	11.4	379	1	Y024_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03
581	4	11.4	294	1	Y301_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	654	4	11.4	379	1	YAD2_YEAST	HYPOTHETICAL 42.5 KD P	2.67e+03
582	4	11.4	295	1	YDH2_XANAU	HYPOTHETICAL PROTEIN I	2.67e+03	655	4	11.4	380	1	Y4CF_RHISN	HYPOTHETICAL 41.8 KD P	2.67e+03
583	4	11.4	296	1	Y1DJ_MYCTU	HYPOTHETICAL 33.8 KD P	2.67e+03	656	4	11.4	382	1	YMF4_CAEEL	HYPOTHETICAL 44.3 KD P	2.67e+03
584	4	11.4	297	1	Y315_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	657	4	11.4	382	1	YMF4_CAEEL	HYPOTHETICAL 43.1 KD P	2.67e+03
585	4	11.4	297	1	YH45_YEAST	HYPOTHETICAL 33.6 KD P	2.67e+03	658	4	11.4	384	1	YXLR_BACSU	XYLOSE REPRESSOR	2.67e+03
586	4	11.4	297	1	YTM4_YEAST	HYPOTHETICAL OXIDOREDU	2.67e+03	659	4	11.4	385	1	YX02_CAEEL	HYPOTHETICAL 39.8 KD P	2.67e+03
587	4	11.4	298	1	YH05_YEAST	HYPOTHETICAL 34.5 KD P	2.67e+03	660	4	11.4	385	1	YX01_MYCTU	HYPOTHETICAL 39.8 KD P	2.67e+03
588	4	11.4	299	1	YNP6_YEAST	HYPOTHETICAL 33.7 KD P	2.67e+03	661	4	11.4	385	1	YX01_MYCTU	HOMEOBOX PROTEIN YX01	2.67e+03
589	4	11.4	300	1	Y43E_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03	662	4	11.4	388	1	YX01_ECOLI	HYPOTHETICAL 43.4 KD P	2.67e+03
590	4	11.4	301	1	Y186_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03	663	4	11.4	388	1	YH09_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
591	4	11.4	308	1	Y209_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	664	4	11.4	389	1	YMF7_CAEEL	HYPOTHETICAL 45.1 KD P	2.67e+03
592	4	11.4	308	1	YB1N_ECOLI	HYPOTHETICAL 34.2 KD P	2.67e+03	665	4	11.4	391	1	Y025_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
593	4	11.4	309	1	Y222_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	666	4	11.4	392	1	YQF6_CAEEL	HYPOTHETICAL 45.3 KD P	2.67e+03
594	4	11.4	309	1	YNR7_YEAST	HYPOTHETICAL 34.9 KD P	2.67e+03	667	4	11.4	392	1	YQ56_CAEEL	HYPOTHETICAL 45.1 KD P	2.67e+03
595	4	11.4	309	1	YOH1_AZOVI	HYPOTHETICAL 33.2 KD P	2.67e+03	668	4	11.4	395	1	YX15_YEAST	HYPOTHETICAL 44.9 KD P	2.67e+03
596	4	11.4	310	1	YH01_ECOLI	HYPOTHETICAL 34.8 KD P	2.67e+03	669	4	11.4	396	1	Y111_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
597	4	11.4	310	1	YB99_SCHPO	HYPOTHETICAL 35.0 KD P	2.67e+03	670	4	11.4	397	1	YMB1_CAEEL	HYPOTHETICAL 44.5 KD P	2.67e+03
598	4	11.4	310	1	YOH1_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	671	4	11.4	399	1	Y874_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
599	4	11.4	317	1	YX32_PYRHO	HYPOTHETICAL PROTEIN P	2.67e+03	672	4	11.4	400	1	YX38_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
600	4	11.4	317	1	Y511_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03	673	4	11.4	400	1	YX23_YEAST	HYPOTHETICAL 45.9 KD P	2.67e+03
601	4	11.4	318	1	YKTB_CAEEL	HYPOTHETICAL 36.1 KD P	2.67e+03	674	4	11.4	401	1	YX07_CAEEL	HYPOTHETICAL 46.7 KD P	2.67e+03
602	4	11.4	319	1	YH41_CRVPA	HYPOTHETICAL PROTEIN I	2.67e+03	675	4	11.4	402	1	YIN2_STRAM	HYPOTHETICAL 44.6 KD P	2.67e+03
603	4	11.4	319	1	YK01_CAEEL	HYPOTHETICAL 35.5 KD P	2.67e+03	676	4	11.4	403	1	YDE4_SCHPO	HYPOTHETICAL 46.5 KD P	2.67e+03
604	4	11.4	320	1	Y054_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03	677	4	11.4	404	1	YX8B_YEAST	HYPOTHETICAL 47.1 KD P	2.67e+03
605	4	11.4	320	1	YB9K_YEAST	HYPOTHETICAL 36.0 KD P	2.67e+03	678	4	11.4	404	1	Y432_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03
606	4	11.4	321	1	YABC_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	679	4	11.4	405	1	Y4RN_RHISN	HYPOTHETICAL 41.6 KD P	2.67e+03
607	4	11.4	322	1	Y348_MYCGE	HYPOTHETICAL LIPOPROTE	2.67e+03	680	4	11.4	405	1	YCC8_YEAST	HYPOTHETICAL 42.6 KD P	2.67e+03

681	4	11.4	4	11.4	1	YG43_YEAST	HYPOTHETICAL 47.3 KD P	2.67e+03	505	1	YLIR_YEAST	HYPOTHETICAL 58.3 KD P	2.67e+03
682	4	11.4	4	11.4	1	Y4DM_RHISN	HYPOTHETICAL 45.3 KD P	2.67e+03	505	1	Y038_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03
683	4	11.4	4	11.4	1	Y148_MVCGE	HYPOTHETICAL PROTEIN M	2.67e+03	507	1	YG2M_YEAST	HYPOTHETICAL 58.2 KD P	2.67e+03
684	4	11.4	4	11.4	1	YIHS_ECOLI	HYPOTHETICAL 47.4 KD P	2.67e+03	507	1	YK4_YEAST	HYPOTHETICAL 59.2 KD P	2.67e+03
685	4	11.4	4	11.4	1	YP71_MYCMI	HYPOTHETICAL PROTEIN I	2.67e+03	508	1	YF00_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
686	4	11.4	4	11.4	1	Y878_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	511	1	YNJC_ECOLI	HYPOTHETICAL 56.2 KD P	2.67e+03
687	4	11.4	4	11.4	1	YS25_CAEEL	HYPOTHETICAL 45.3 KD P	2.67e+03	511	1	YDEX_ECOLI	HYPOTHETICAL ABC TRANS	2.67e+03
688	4	11.4	4	11.4	1	YCHO_ECOLI	HYPOTHETICAL 47.0 KD P	2.67e+03	512	1	Y994_SCHPO	PROBABLE ZINC METALLOP	2.67e+03
689	4	11.4	4	11.4	1	Y943_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	516	1	YVNF_AZOCH	HYPOTHETICAL PROTEIN I	2.67e+03
690	4	11.4	4	11.4	1	YG44_YEAST	HYPOTHETICAL 46.4 KD T	2.67e+03	516	1	YIJ7_YEAST	HYPOTHETICAL 59.9 KD P	2.67e+03
691	4	11.4	4	11.4	1	YS48_TREVE	SATELLITE RNA 48 KD PR	2.67e+03	517	1	XYLB_BUTFI	BETA-XYLOSIDASE (EC 3.	2.67e+03
692	4	11.4	4	11.4	1	YBIU_ECOLI	HYPOTHETICAL 47.3 KD P	2.67e+03	518	1	YB95_METJA	PUTATIVE 2-ISOPROPYLMA	2.67e+03
693	4	11.4	4	11.4	1	YDW4_SCHPO	HYPOTHETICAL 48.5 KD P	2.67e+03	524	1	Y4LR_RHISN	HYPOTHETICAL 56.8 KD P	2.67e+03
694	4	11.4	4	11.4	1	ZP3_MESAU	ZONA PELLUCIDA SPERM-B	2.67e+03	524	1	YMS4_YEAST	HYPOTHETICAL 60.5 KD P	2.67e+03
695	4	11.4	4	11.4	1	Y46K_DROME	HYPOTHETICAL 46.0 KD P	2.67e+03	525	1	YB8_YEAST	HYPOTHETICAL 60.1 KD P	2.67e+03
696	4	11.4	4	11.4	1	ZP3_FELUCA	ZONA PELLUCIDA SPERM-B	2.67e+03	528	1	YLD4_CAEEL	HYPOTHETICAL 57.3 KD P	2.67e+03
697	4	11.4	4	11.4	1	ZP3_MOUSE	ZONA PELLUCIDA SPERM-B	2.67e+03	530	1	Y083_MYCTU	HYPOTHETICAL 58.3 KD P	2.67e+03
698	4	11.4	4	11.4	1	ZP3_CANFA	ZONA PELLUCIDA SPERM-B	2.67e+03	534	1	YLA9_SCHPO	HYPOTHETICAL 54.2 KD S	2.67e+03
699	4	11.4	4	11.4	1	WCAK_ECOLI	COLANIC ACID BIOSYNTH	2.67e+03	534	1	YE25_METTH	O-STALOGLYCOPROTEIN EN	2.67e+03
700	4	11.4	4	11.4	1	YEBS_ECOLI	HYPOTHETICAL 48.3 KD P	2.67e+03	537	1	YAJE_SCHPO	HYPOTHETICAL 59.0 KD P	2.67e+03
701	4	11.4	4	11.4	1	YFJD_ECOLI	HYPOTHETICAL 46.4 KD P	2.67e+03	537	1	YIV9_YEAST	POTATIVE ASPARYL PROT	2.67e+03
702	4	11.4	4	11.4	1	Y831_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	538	1	YIEH_SCHPO	POTENTIAL OUTER MEMBRA	2.67e+03
703	4	11.4	4	11.4	1	Y103_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	539	1	YIEC_ECOLI	HYPOTHETICAL 62.6 KD P	2.67e+03
704	4	11.4	4	11.4	1	YQD7_CAEEL	HYPOTHETICAL 49.2 KD P	2.67e+03	543	1	YLLA_BACSU	ZINC FINGER PROTEIN 8	2.67e+03
705	4	11.4	4	11.4	1	YIN5_YEAST	HYPOTHETICAL 48.0 KD P	2.67e+03	545	1	YB90_YEAST	HYPOTHETICAL 61.3 KD P	2.67e+03
706	4	11.4	4	11.4	1	YHG1_YEAST	HYPOTHETICAL 49.8 KD P	2.67e+03	547	1	YEE0_ECOLI	HYPOTHETICAL 60.1 KD P	2.67e+03
707	4	11.4	4	11.4	1	Y021_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03	550	1	Y436_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
708	4	11.4	4	11.4	1	YDW1_SCHPO	HYPOTHETICAL 49.2 KD P	2.67e+03	552	1	Y4HP_RHISN	HYPOTHETICAL 61.7 KD P	2.67e+03
709	4	11.4	4	11.4	1	YAV4_SCHPO	HYPOTHETICAL 51.9 KD P	2.67e+03	554	1	YF61_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
710	4	11.4	4	11.4	1	YIHI_YEAST	HYPOTHETICAL 51.3 KD P	2.67e+03	557	1	YG62_ARATH	HYPOTHETICAL 61.8 KD T	2.67e+03
711	4	11.4	4	11.4	1	YOR3_GLUUS	HYPOTHETICAL PROTEIN I	2.67e+03	559	1	Y289_HUMAN	HYPOTHETICAL PROLINE-R	2.67e+03
712	4	11.4	4	11.4	1	Y4WB_RHISN	HYPOTHETICAL ZINC PROT	2.67e+03	561	1	YGG4_YEAST	POTATIVE ATP-DEPENDENT	2.67e+03
713	4	11.4	4	11.4	1	YJ94_YEAST	HYPOTHETICAL 49.7 KD P	2.67e+03	566	1	YQK4_CAEEL	ASPARTIC PROTEINASE 3	2.67e+03
714	4	11.4	4	11.4	1	YPEB_BACSU	HYPOTHETICAL 51.2 KD P	2.67e+03	571	1	YAP3_YEAST	ALPHA-L-ARABINOFURANOS	2.67e+03
715	4	11.4	4	11.4	1	YJ66_YEAST	HYPOTHETICAL 52.1 KD P	2.67e+03	572	1	YHM4_YEAST	HEAT SHOCK PROTEIN 70	2.67e+03
716	4	11.4	4	11.4	1	VU10_HSV7J	U10 PROTEIN.	2.67e+03	574	1	YJ23_YEAST	HYPOTHETICAL 66.1 KD P	2.67e+03
717	4	11.4	4	11.4	1	Y003_BORBU	HYPOTHETICAL PROTEIN B	2.67e+03	576	1	YIT0_YEAST	HYPOTHETICAL 65.7 KD P	2.67e+03
718	4	11.4	4	11.4	1	YM18_YEAST	HYPOTHETICAL 51.7 KD P	2.67e+03	578	1	YAG3_SCHPO	HYPOTHETICAL 62.1 KD P	2.67e+03
719	4	11.4	4	11.4	1	YHCL_ECOLI	HYPOTHETICAL 48.8 KD P	2.67e+03	580	1	YDEM_CAEEL	HYPOTHETICAL 65.5 KD T	2.67e+03
720	4	11.4	4	11.4	1	YP95_YEAST	HYPOTHETICAL 51.7 KD P	2.67e+03	580	1	YHA2_SIKCO	HYPOTHETICAL 66.3 KD P	2.67e+03
721	4	11.4	4	11.4	1	YC47_SYNY3	HYPOTHETICAL 52.7 KD P	2.67e+03	583	1	Y0A3_MYCTU	HYPOTHETICAL 62.1 KD P	2.67e+03
722	4	11.4	4	11.4	1	YDGF_BACSU	HYPOTHETICAL TRANSPORT	2.67e+03	585	1	YB3M_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
723	4	11.4	4	11.4	1	YNE1_CAEEL	HYPOTHETICAL 52.8 KD P	2.67e+03	586	1	YG2B_YEAST	HYPOTHETICAL 65.5 KD P	2.67e+03
724	4	11.4	4	11.4	1	YTPL_YEAST	YTPL PROTEIN.	2.67e+03	590	1	YMW2_YEAST	HYPOTHETICAL 67.6 KD P	2.67e+03
725	4	11.4	4	11.4	1	YEBD_SCHPO	HYPOTHETICAL 53.5 KD P	2.67e+03	590	1	YKTA_CAEEL	HYPOTHETICAL 65.5 KD P	2.67e+03
726	4	11.4	4	11.4	1	YFHD_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03	593	1	YHYA_BPH44	HYPOTHETICAL 65 KD PRO	2.67e+03
727	4	11.4	4	11.4	1	Y123_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03	594	1	YKA5_YEAST	HYPOTHETICAL 67.9 KD P	2.67e+03
728	4	11.4	4	11.4	1	YA65_MOUSE	65 KD YES-ASSOCIATED P	2.67e+03	598	1	YAU5_SCHPO	HYPOTHETICAL 60.7 KD P	2.67e+03
729	4	11.4	4	11.4	1	YVDB_MOUSE	VITAMIN D-BINDING PROT	2.67e+03	599	1	YGM5_YEAST	HYPOTHETICAL 68.5 KD P	2.67e+03
730	4	11.4	4	11.4	1	YVDB_ECOLI	PUTATIVE BETAIN ALDEH	2.67e+03	608	1	YD6C_SCHPO	HYPOTHETICAL 67.5 KD P	2.67e+03
731	4	11.4	4	11.4	1	YVDB_HUMAN	VITAMIN D-BINDING PROT	2.67e+03	609	1	YSW1_YEAST	SPORE-SPECIFIC PROTEIN	2.67e+03
732	4	11.4	4	11.4	1	Y123_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03	612	1	YND4_YEAST	HYPOTHETICAL 69.4 KD P	2.67e+03
733	4	11.4	4	11.4	1	YIGN_ECOLI	HYPOTHETICAL 54.7 KD P	2.67e+03	614	1	YB83_YEAST	HYPOTHETICAL 69.6 KD P	2.67e+03
734	4	11.4	4	11.4	1	YVDB_RAT	VITAMIN D-BINDING PROT	2.67e+03	614	1	Y383_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
735	4	11.4	4	11.4	1	YVDB_RABIT	VITAMIN D-BINDING PROT	2.67e+03	616	1	Y3A3_SCHPO	HYPOTHETICAL 67.3 KD T	2.67e+03
736	4	11.4	4	11.4	1	Y481_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03	620	1	Y241_MYCGE	HYPOTHETICAL 68.2 KD P	2.67e+03
737	4	11.4	4	11.4	1	Y142_NVPDP	HYPOTHETICAL 55.8 KD P	2.67e+03	621	1	YQ0A_CAEEL	HYPOTHETICAL 69.5 KD P	2.67e+03
738	4	11.4	4	11.4	1	YAV5_SCHPO	PROBABLE AMINOTRANSFER	2.67e+03	621	1	YQ01_CAEEL	HYPOTHETICAL 69.9 KD P	2.67e+03
739	4	11.4	4	11.4	1	WAC_BP74	FIBRININ (WHISKER ANTI	2.67e+03	622	1	YAE7_YEAST	HYPOTHETICAL 72.1 KD P	2.67e+03
740	4	11.4	4	11.4	1	YU2_CAEEL	HYPOTHETICAL 54.5 KD P	2.67e+03	623	1	Y014_MYCGE	HYPOTHETICAL ABC TRANS	2.67e+03
741	4	11.4	4	11.4	1	YDP8_SCHPO	HYPOTHETICAL 54.2 KD T	2.67e+03	630	1	Y242_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
742	4	11.4	4	11.4	1	YD3C_YEAST	HYPOTHETICAL 55.2 KD P	2.67e+03	632	1	Y242_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03
743	4	11.4	4	11.4	1	YE01_SCHPO	HYPOTHETICAL 54.3 KD T	2.67e+03	632	1	YKGS_CAEEL	HYPOTHETICAL 71.9 KD P	2.67e+03
744	4	11.4	4	11.4	1	YF86_METJA	HYPOTHETICAL PROTEIN M	2.67e+03	641	1	YGG0_YEAST	HYPOTHETICAL 73.1 KD P	2.67e+03
745	4	11.4	4	11.4	1	YAZ9_SCHPO	HYPOTHETICAL 55.1 KD P	2.67e+03	642	1	YB9W_YEAST	HYPOTHETICAL 71.1 KD P	2.67e+03
746	4	11.4	4	11.4	1	YNM4_YEAST	HYPOTHETICAL 54.9 KD P	2.67e+03					
747	4	11.4	4	11.4	1	YGFH_ECOLI	HYPOTHETICAL 53.8 KD P	2.67e+03					
748	4	11.4	4	11.4	1	YD48_SCHPO	HYPOTHETICAL 56.5 KD P	2.67e+03					
749	4	11.4	4	11.4	1	YE11_YEAST	55.2 KD PROTEIN IN HXT	2.67e+03					
750	4	11.4	4	11.4	1	YK01_CAEEL	HYPOTHETICAL 55.9 KD P	2.67e+03					
751	4	11.4	4	11.4	1	YP25_CAEEL	HYPOTHETICAL 56.2 KD P	2.67e+03					
752	4	11.4	4	11.4	1	YAO1_SCHPO	POTATIVE TRANSPORTER C	2.67e+03					
753	4	11.4	4	11.4	1	YDL1_SCHPO	HYPOTHETICAL 57.9 KD P	2.67e+03					

827	4	11.4	4	11.4	644	1	YED0_YEAST	HEAT SHOCK PROTEIN 70	2.67e+03
828	4	11.4	4	11.4	646	1	YHDA_ECOLI	HYPOTHETICAL 73.3 KD P	2.67e+03
829	4	11.4	4	11.4	646	1	WEEL1_HUMAN	WEEL1-LIKE PROTEIN KINA	2.67e+03
830	4	11.4	4	11.4	646	1	WEEL1_MOUSE	WEEL1-LIKE PROTEIN KINA	2.67e+03
831	4	11.4	4	11.4	653	1	Y07S_MYCTU	HYPOTHETICAL 69.5 KD P	2.67e+03
832	4	11.4	4	11.4	653	1	YKR2_CAEEL	HYPOTHETICAL 70.9 KD P	2.67e+03
833	4	11.4	4	11.4	655	1	YJ68_YEAST	HYPOTHETICAL 74.1 KD P	2.67e+03
834	4	11.4	4	11.4	655	1	YDH2_SCHPO	HYPOTHETICAL 75.4 KD P	2.67e+03
835	4	11.4	4	11.4	655	1	YJHG_ECOLI	HYPOTHETICAL 70.0 KD P	2.67e+03
836	4	11.4	4	11.4	656	1	YMD4_YEAST	HYPOTHETICAL 74.2 KD P	2.67e+03
837	4	11.4	4	11.4	659	1	YK16_YEAST	HYPOTHETICAL 74.7 KD T	2.67e+03
838	4	11.4	4	11.4	659	1	Y03L_MYCLE	HYPOTHETICAL 71.2 KD P	2.67e+03
839	4	11.4	4	11.4	659	1	YJET_BACSU	HYPOTHETICAL 74.3 KD P	2.67e+03
840	4	11.4	4	11.4	659	1	YH5_YEAST	HYPOTHETICAL 75.9 KD P	2.67e+03
841	4	11.4	4	11.4	666	1	YEA7_YEAST	HYPOTHETICAL 72.5 KD P	2.67e+03
842	4	11.4	4	11.4	666	1	Y079_SCHPO	HYPOTHETICAL 74.2 KD P	2.67e+03
843	4	11.4	4	11.4	675	1	VP55_YEAST	VACUOLAR PROTEIN SORTI	2.67e+03
844	4	11.4	4	11.4	681	1	YADD_SCHPO	HYPOTHETICAL 74.2 KD P	2.67e+03
845	4	11.4	4	11.4	683	1	YJCO_ECOLI	HYPOTHETICAL 76.1 KD P	2.67e+03
846	4	11.4	4	11.4	686	1	YPR4_CAEEL	HYPOTHETICAL 77.0 KD T	2.67e+03
847	4	11.4	4	11.4	687	1	YB50_YEAST	HYPOTHETICAL 78.8 KD P	2.67e+03
848	4	11.4	4	11.4	688	1	Y07_YEAST	HYPOTHETICAL 79.5 KD P	2.67e+03
849	4	11.4	4	11.4	688	1	YB9F_YEAST	HYPOTHETICAL 80.4 KD P	2.67e+03
850	4	11.4	4	11.4	690	1	YNP9_CAEEL	HYPOTHETICAL 79.2 KD P	2.67e+03
851	4	11.4	4	11.4	693	1	YETI_SCHPO	PUTATIVE ABC TRANSPORT	2.67e+03
852	4	11.4	4	11.4	698	1	YMCA_ECOLI	HYPOTHETICAL 78.7 KD L	2.67e+03
853	4	11.4	4	11.4	698	1	YB06_YEAST	HYPOTHETICAL 79.2 KD P	2.67e+03
854	4	11.4	4	11.4	701	1	YSS1_CAEEL	HYPOTHETICAL 80.3 KD T	2.67e+03
855	4	11.4	4	11.4	701	1	YAB5_SCHPO	HYPOTHETICAL 81.8 KD P	2.67e+03
856	4	11.4	4	11.4	702	1	YAI4_YEAST	HYPOTHETICAL 78.3 KD P	2.67e+03
857	4	11.4	4	11.4	703	1	Y411_RHISN	HYPOTHETICAL 76.2 KD P	2.67e+03
858	4	11.4	4	11.4	704	1	YJ93_YEAST	HYPOTHETICAL 82.8 KD P	2.67e+03
859	4	11.4	4	11.4	707	1	YJ9C_YEAST	HYPOTHETICAL 81.2 KD P	2.67e+03
860	4	11.4	4	11.4	722	1	Y022_TREPA	HYPOTHETICAL PROTEIN T	2.67e+03
861	4	11.4	4	11.4	726	1	YANA_RHISN	PROBABLE PEPTIDASE Y4N	2.67e+03
862	4	11.4	4	11.4	732	1	YAGR_ECOLI	HYPOTHETICAL 78.1 KD P	2.67e+03
863	4	11.4	4	11.4	732	1	YMM1_CAEEL	HYPOTHETICAL 81.8 KD P	2.67e+03
864	4	11.4	4	11.4	734	1	YD2B_SCHPO	HYPOTHETICAL 85.4 KD P	2.67e+03
865	4	11.4	4	11.4	734	1	Y04H_MYCTU	HYPOTHETICAL 72.9 KD P	2.67e+03
866	4	11.4	4	11.4	740	1	YHGF_ECOLI	HYPOTHETICAL 81.4 KD P	2.67e+03
867	4	11.4	4	11.4	745	1	YJ00_YEAST	HYPOTHETICAL 84.3 KD P	2.67e+03
868	4	11.4	4	11.4	750	1	YF03_MYCPN	HYPOTHETICAL 85.3 KD P	2.67e+03
869	4	11.4	4	11.4	751	1	Y236_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
870	4	11.4	4	11.4	753	1	YJ05_CAEEL	HYPOTHETICAL 85.1 KD N	2.67e+03
871	4	11.4	4	11.4	753	1	YPGA_CAEEL	HYPOTHETICAL 86.5 KD P	2.67e+03
872	4	11.4	4	11.4	756	1	Y328_MYCGE	HYPOTHETICAL PROTEIN M	2.67e+03
873	4	11.4	4	11.4	759	1	Y353_METJA	HYPOTHETICAL MCM-TYPE	2.67e+03
874	4	11.4	4	11.4	760	1	YCE5_YEAST	HYPOTHETICAL 87.2 KD P	2.67e+03
875	4	11.4	4	11.4	760	1	YAD1_YEAST	HYPOTHETICAL 86.6 KD P	2.67e+03
876	4	11.4	4	11.4	762	1	Y951_METJA	HYPOTHETICAL MCM-TYPE	2.67e+03
877	4	11.4	4	11.4	765	1	Y260_MYCGE	HYPOTHETICAL PROTEIN K	2.67e+03
878	4	11.4	4	11.4	765	1	Y008_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
879	4	11.4	4	11.4	775	1	YPA_NCDV	OUTER CAPSID PROTEIN V	2.67e+03
880	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
881	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
882	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
883	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
884	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
885	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
886	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
887	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
888	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
889	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
890	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
891	4	11.4	4	11.4	776	1	YPA_ROTIRH	OUTER CAPSID PROTEIN V	2.67e+03
892	4	11.4	4	11.4	779	1	YHUN_ECOLI	HYPOTHETICAL 86.0 KD P	2.67e+03
893	4	11.4	4	11.4	782	1	XPB_HUMAN	DNA-REPAIR PROTEIN COM	2.67e+03
894	4	11.4	4	11.4	786	1	YMM1_CAEEL	HYPOTHETICAL 88.1 KD P	2.67e+03
895	4	11.4	4	11.4	798	1	YLFN_CAEEL	HYPOTHETICAL 89.8 KD P	2.67e+03
896	4	11.4	4	11.4	799	1	ZY11_CAEEL	EARLY EMBRYOGENESIS 2Y	2.67e+03
897	4	11.4	4	11.4	807	1	YK59_YEAST	HYPOTHETICAL 91.7 KD T	2.67e+03
898	4	11.4	4	11.4	838	1	YMS9_YEAST	HYPOTHETICAL 96.8 KD P	2.67e+03
899	4	11.4	4	11.4	838	1	YMC9_YEAST	HYPOTHETICAL 96.7 KD P	2.67e+03

900	4	11.4	4	11.4	843	1	Y096_CAEEL	HYPOTHETICAL 93.9 KD P	2.67e+03
901	4	11.4	4	11.4	849	1	Y343_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
902	4	11.4	4	11.4	871	1	YAD7_SCHPO	HYPOTHETICAL 98.4 KD P	2.67e+03
903	4	11.4	4	11.4	875	1	Y066_NPVP	HYPOTHETICAL 98.6 KD P	2.67e+03
904	4	11.4	4	11.4	878	1	YB9X_YEAST	HYPOTHETICAL 98.1 KD T	2.67e+03
905	4	11.4	4	11.4	881	1	YU78_YEAST	HYPOTHETICAL 97.7 KD M	2.67e+03
906	4	11.4	4	11.4	881	1	YU78_HAEIN	HYPOTHETICAL PROTEIN H	2.67e+03
907	4	11.4	4	11.4	882	1	YBAH_SCHPO	HYPOTHETICAL 99.0 KD P	2.67e+03
908	4	11.4	4	11.4	882	1	YK1_CAEEL	HYPOTHETICAL 98.3 KD P	2.67e+03
909	4	11.4	4	11.4	883	1	YHL6_YEAST	PUTATIVE 101.3 KD TRAN	2.67e+03
910	4	11.4	4	11.4	885	1	Y143_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
911	4	11.4	4	11.4	886	1	YB86_SCHPO	HYPOTHETICAL 98.3 KD P	2.67e+03
912	4	11.4	4	11.4	886	1	YUHO_YEAST	HYPOTHETICAL 104.3 KD	2.67e+03
913	4	11.4	4	11.4	892	1	YU06_YEAST	HYPOTHETICAL 103.5 KD	2.67e+03
914	4	11.4	4	11.4	907	1	Y274_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
915	4	11.4	4	11.4	924	1	YB53_YEAST	HYPOTHETICAL 104.5 KD	2.67e+03
916	4	11.4	4	11.4	925	1	YBK4_YEAST	HYPOTHETICAL 104.7 KD	2.67e+03
917	4	11.4	4	11.4	925	1	YPR1_BPHPI	PROBABLE TAIL FIBRE PR	2.67e+03
918	4	11.4	4	11.4	928	1	YU23_YEAST	HYPOTHETICAL 106.7 KD	2.67e+03
919	4	11.4	4	11.4	929	1	YDM6_SCHPO	HYPOTHETICAL 105.1 KD	2.67e+03
920	4	11.4	4	11.4	932	1	YMB4_CAEEL	HYPOTHETICAL 104.4 KD	2.67e+03
921	4	11.4	4	11.4	934	1	Y321_MYCGE	HYPOTHETICAL LIPOPROTE	2.67e+03
922	4	11.4	4	11.4	935	1	YAC6_SCHPO	HYPOTHETICAL 105.6 KD	2.67e+03
923	4	11.4	4	11.4	943	1	YLM5_CAEEL	HYPOTHETICAL 105.9 KD	2.67e+03
924	4	11.4	4	11.4	952	1	YK15_CAEEL	HYPOTHETICAL 105.3 KD	2.67e+03
925	4	11.4	4	11.4	954	1	YU59_YEAST	HYPOTHETICAL 108.7 KD	2.67e+03
926	4	11.4	4	11.4	960	1	YMX6_YEAST	HYPOTHETICAL 105.9 KD	2.67e+03
927	4	11.4	4	11.4	964	1	YQY1_CAEEL	HYPOTHETICAL 108.9 KD	2.67e+03
928	4	11.4	4	11.4	971	1	YQIA_CAEEL	HYPOTHETICAL 111.8 KD	2.67e+03
929	4	11.4	4	11.4	981	1	YU48_YEAST	HYPOTHETICAL 113.2 KD	2.67e+03
930	4	11.4	4	11.4	983	1	Y144_HUMAN	HYPOTHETICAL PROTEIN K	2.67e+03
931	4	11.4	4	11.4	996	1	YHN3_YEAST	HYPOTHETICAL 113.8 KD	2.67e+03
932	4	11.4	4	11.4	997	1	Y414_MYCPN	HYPOTHETICAL PROTEIN M	2.67e+03
933	4	11.4	4	11.4	997	1	YEBG_SCHPO	HYPOTHETICAL 112.6 KD	2.67e+03
934	4	11.4	4	11.4	1002	1	YEMA_DROME	YEMANUCLEIN-ALPHA	2.67e+03
935	4	11.4	4	11.4	1004	1	YU21_YEAST	HYPOTHETICAL 113.9 KD	2.67e+03
936	4	11.4	4	11.4	1005	1	YCF1_OENBE	HYPOTHETICAL PROTEIN (	2.67e+03
937	4	11.4	4	11.4	1005	1	Y321_MYCPN	HYPOTHETICAL LIPOPROTE	2.67e+03
938	4	11.4	4	11.4	1010	1	Y651_HAEIN	PROBABLE TONB-DEPEND	2.67e+03
939	4	11.4	4	11.4	1010	1	YNT5_DROME	PROTEIN DWNT-5 PRECURS	2.67e+03
940	4	11.4	4	11.4	1018	1	YC14_METJA	HYPOTHETICAL PROTEIN M	2.67e+03
941	4	11.4	4	11.4	1021	1	YF77_CAEEL	HYPOTHETICAL 111.7 KD	2.67e+03
942	4	11.4	4	11.4	1030	1	YEO2_SCHPO	PUTATIVE HELICASE CLTH	2.67e+03
943	4	11.4	4	11.4	1033	1	YD55_SCHPO	PROBABLE CATION-TRANSP	2.67e+03
944	4	11.4	4	11.4	1040	1	YEGN_ECOLI	HYPOTHETICAL 112.1 KD	2.67e+03
945	4	11.4	4	11.4	1044	1	YDE1_SCHPO	HYPOTHETICAL PROTEIN C	2.67e+03
946	4	11.4	4	11.4	1056	1	YNN2_YEAST	HYPOTHETICAL 119.3 KD	2.67e+03
947	4	11.4	4	11.4	1060	1	YNN8_YEAST	HYPOTHETICAL 118.3 KD	2.67e+03
948	4	11.4	4	11.4	1068	1	YCF0_MARPO	HYPOTHETICAL 127 KD PR	2.67e+03
949	4	11.4	4	11.4	1088	1	YMS4_YEAST	HYPOTHETICAL 126.6 KD	2.67e+03
950	4	11.4	4	11.4	1102	1	YM66_YEAST	HYPOTHETICAL 128.1 KD	2.67e+03
951	4	11.4	4	11.4	1113	1	Y140_MYCPN	HYPOTHETICAL ATP-BINDI	2.67e+03
952	4	11.4	4	11.4	1121	1	YU09_YEAST	HYPOTHETICAL 127.4 KD	2.67e+03
953	4	11.4	4	11.4	1130	1	YU17_CAEEL	HYPOTHETICAL 131.5 KD	2.67e+03
954	4	11.4	4	11.4	1154	1	WCL1_NEUCR	WHITE COLLAR 1 PROTEIN	2.67e+03

973 4 11.4 1356 1 Y195\_HUMAN HYPOTHETICAL PROTEIN K 2.67e+03  
 974 4 11.4 1375 1 Y1P9\_YEAST HYPOTHETICAL 156.9 KD 2.67e+03  
 975 4 11.4 1385 1 YMS5\_CAEEL HYPOTHETICAL 159.2 KD 2.67e+03  
 976 4 11.4 1386 1 Y064\_MYCPN HYPOTHETICAL PROTEIN M 2.67e+03  
 977 4 11.4 1401 1 WRN\_MOUSE WERNER SYNDROME HELICA 2.67e+03  
 978 4 11.4 1403 1 YDF3\_SCHPO PROBABLE EUKARYOTIC IN 2.67e+03  
 979 4 11.4 1416 1 YN81\_CAEEL HYPOTHETICAL 316.1 KD 2.67e+03  
 980 4 11.4 1435 1 Y194\_HUMAN HYPOTHETICAL PROTEIN K 2.67e+03  
 981 4 11.4 1562 1 YM81\_YEAST HYPOTHETICAL 180.2 KD 2.67e+03  
 982 4 11.4 1569 1 YPJA\_ECOLI HYPOTHETICAL 98.4 KD P 2.67e+03  
 983 4 11.4 1583 1 YA45\_SCHPO HYPOTHETICAL 180.2 KD 2.67e+03  
 984 4 11.4 1584 1 YJ9G\_YEAST HYPOTHETICAL 182.0 KD 2.67e+03  
 985 4 11.4 1658 1 YM67\_YEAST HYPOTHETICAL 187.1 KD 2.67e+03  
 986 4 11.4 1733 1 VNUA\_PRYKA PROBABLE NUCLEAR ANTIG 2.67e+03  
 987 4 11.4 1750 1 Y832\_METJA HYPOTHETICAL PROTEIN M 2.67e+03  
 988 4 11.4 1756 1 YCF1\_PINTH HYPOTHETICAL 205.3 KD 2.67e+03  
 989 4 11.4 1769 1 YJK9\_YEAST HYPOTHETICAL 200.0 KD 2.67e+03  
 990 4 11.4 1804 1 YFA7\_YEAST HYPOTHETICAL 207.6 KD 2.67e+03  
 991 4 11.4 1822 1 YM68\_CAEEL HYPOTHETICAL HELICASE 2.67e+03  
 992 4 11.4 1984 1 YLDR0ME PUTATIVE VITELLOGENIN 2.67e+03  
 993 4 11.4 2052 1 YDOB\_SCHPO HYPOTHETICAL 234.0 KD 2.67e+03  
 994 4 11.4 2054 1 YCF2\_PINTH HYPOTHETICAL 244.6 KD 2.67e+03  
 995 4 11.4 2136 1 YCF2\_MARPO HYPOTHETICAL 259 KD PR 2.67e+03  
 996 4 11.4 2216 1 YCF2\_EPIVI HYPOTHETICAL 260 KD PR 2.67e+03  
 997 4 11.4 2670 1 YAO5\_SCHPO PUTATIVE TRANSLATIONAL 2.67e+03  
 998 4 11.4 3144 1 YP13\_YEAST VACUOLAR PROTEIN SORTI 2.67e+03  
 999 4 11.4 3343 1 YOG7\_CAEEL HYPOTHETICAL 375.7 KD 2.67e+03  
 1000 4 11.4 4385 1 YP73\_CAEEL HYPOTHETICAL 486.0 KD 2.67e+03

## ALIGNMENTS

RESULT 1  
 ID SCP2\_BRALA STANDARD; PRT: 185 AA.  
 AC P04570;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE SARCOPLASMIC CALCIUM-BINDING PROTEINS II, V, VI, AND AND VII (SCP II, V, VI, VII).  
 DE BRANCHIOSTOMA LANCEOLATUM (COMMON LANCELET) (AMPHIOXUS).  
 OC EUKARYOTA; METAZOA; CHORDATA; CEPHALOCHORDATA; BRANCHIOSTOMIDAE;  
 OC BRANCHIOSTOMA.  
 RN [1]  
 RP SEQUENCE.  
 RA TAKAGI T., KONISHI K., COX J.A.;  
 RT "Amino acid sequence of two sarcoplasmic calcium-binding proteins from the protochordate amphioxus."  
 RL BIOCHEMISTRY 25:3585-3592(1986).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE: 91006125.  
 RA TAKAGI T., COX J.A.;  
 RT "Amino acid sequences of four isoforms of amphioxus sarcoplasmic calcium-binding proteins."  
 RL EUR. J. BIOCHEM. 192:387-399(1990).  
 RN [3]  
 RP SEQUENCE (SCP V TO VII).  
 RX MEDLINE: 92339504.  
 RA TAKAGI T., VALETTE-TALBI L., COX J.A.;  
 RT "Primary structure of three minor isoforms of amphioxus sarcoplasmic calcium-binding proteins."  
 RL FEBS LETT. 302:159-160(1992).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE: 93156054.  
 RA COOK W.J., JEFFREY L.C., COX J.A., VIJAY-KUMAR S.;  
 RT "Structure of a sarcoplasmic calcium-binding protein from amphioxus refined at 2.4-A resolution."  
 RL J. MOL. BIOL. 229:461-471(1993).  
 CC -!- FUNCTION: LIKE PARVALBUMINS, SCP'S SEEM TO BE MORE ABUNDANT IN FAST CONTRACTING MUSCLES, BUT NO FUNCTIONAL RELATIONSHIP CAN BE ESTABLISHED FROM THIS DISTRIBUTION.

CC -!- THE SARCOPLASMIC CALCIUM-BINDING PROTEINS ARE ABUNDANT IN THE MUSCLE OF ARTHROPODS, MOLLUSCS, ANNELIDS, AND PROTOCHORDATES.  
 CC -!- THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING SITES;  
 CC -!- POTENTIAL SITE 4 HAS LOST AFFINITY FOR CALCIUM.  
 CC -!- THERE ARE 7 DIFFERENT SCP'S IN AMPHIOXUS. THE SEQUENCE SHOWN HERE IS THAT OF SCP II.  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
 DR PIR: A24479; A24479.  
 DR PIR: S13183; S13183.  
 DR PDB: 2SAS; 31-OCT-93.  
 DR PROSITE: PS00018; EF\_HAND; 3.  
 DR PFAM: PF00036; ehand; 1.  
 KW MUSCLE PROTEIN; CALCIUM-BINDING; DUPLICATION; 3D-STRUCTURE.  
 FT CA\_BIND 19 30 SITE 1.  
 FT CA\_BIND 70 81 SITE 2.  
 FT CA\_BIND 115 126 SITE 3.  
 FT DOMAIN 150 161 ANCESTRAL CALCIUM SITE 4.  
 FT VARIANT 25 25 S -> A (IN SCP V).  
 FT VARIANT 28 28 D -> W (IN SCP V, VI, AND VII).  
 FT VARIANT 29 29 N -> E (IN SCP VI).  
 FT HELIX 4 17  
 FT TURN 18 18  
 FT TURN 20 21  
 FT STRAND 25 26  
 FT HELIX 28 45  
 FT HELIX 48 69  
 FT TURN 71 72  
 FT STRAND 77 78  
 FT HELIX 79 91  
 FT TURN 92 92  
 FT HELIX 96 98  
 FT TURN 101 102  
 FT HELIX 103 114  
 FT TURN 116 117  
 FT HELIX 124 130  
 FT TURN 131 132  
 FT HELIX 141 149  
 FT TURN 150 153  
 FT HELIX 158 170  
 FT HELIX 176 181  
 FT TURN 182 182  
 SQ SEQUENCE 185 AA; 21286 MW; 20716924 CRC32;

Query Match 17.1%; Score 6; DB 1; Length 185;

Best Local Similarity 100.0%; Pred. No. 2.56e+00;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 INKDDV 76

|||||

Qy 16 INKDDV 21

## RESULT 2

ID SCP1\_BRALA STANDARD; PRT: 185 AA.  
 AC P04569;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE SARCOPLASMIC CALCIUM-BINDING PROTEINS I, III, AND IV (SCP I, III, IV).  
 OS BRANCHIOSTOMA LANCEOLATUM (COMMON LANCELET) (AMPHIOXUS).  
 OC EUKARYOTA; METAZOA; CHORDATA; CEPHALOCHORDATA; BRANCHIOSTOMIDAE;  
 OC BRANCHIOSTOMA.  
 RN [1]  
 RP SEQUENCE (SCP I).  
 RA TAKAGI T., KONISHI K., COX J.A.;  
 RT "Amino acid sequence of two sarcoplasmic calcium-binding proteins from the protochordate amphioxus."  
 RL BIOCHEMISTRY 25:3585-3592(1986).  
 RN [2]  
 RP SEQUENCE (SCP III AND IV), AND REVISIONS (SCP I).  
 RX MEDLINE: 91006125.  
 RA TAKAGI T., COX J.A.;  
 RT "Amino acid sequences of four isoforms of amphioxus sarcoplasmic

calcium-binding proteins.";  
EUR. J. BIOCHEM. 192:387-399(1990).  
-!- FUNCTION: LIKE PARVALBUMINS, SCP'S SEEM TO BE MORE ABUNDANT IN  
CC FAST CONTRACTING MUSCLES, BUT NO FUNCTIONAL RELATIONSHIP CAN BE  
CC ESTABLISHED FROM THIS DISTRIBUTION.  
CC -!- THE SARCOPLASMIC CALCIUM-BINDING PROTEINS ARE ABUNDANT IN THE  
CC MUSCLE OF ARTHROPODS, MOLLUSCS, ANNELIDS, AND PROTOCHORDATES.  
CC -!- THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING SITES;  
CC POTENTIAL SITE 4 HAS LOST AFFINITY FOR CALCIUM.  
CC -!- THERE ARE 7 DIFFERENT SCP'S IN AMPHIOXUS. THE SEQUENCE SHOWN  
CC HERE IS THAT OF SCP I.  
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
DR PIR: B24479; B24479.  
DR PIR: S13182; S13182.  
DR PIR: S13184; S13184.  
DR PIR: S13185; S13185.  
DR PROSITE: PS00018; EF\_HAND; 3.  
DR PFAM: PF00036; efhand; 1.  
DR HSP: P04570; 2SAS.  
KW MUSCLE PROTEIN; CALCIUM-BINDING; DUPLICATION.  
FT CA\_BIND 19 30 SITE 1.  
FT CA\_BIND 70 81 SITE 2.  
FT CA\_BIND 115 126 SITE 3.  
FT DOMAIN 150 161 ANCESTRAL CALCIUM SITE 4.  
FT VARIANT 20 20 Y -> M (IN SCP III).  
FT VARIANT 23 23 D -> N (IN SCP IV).  
SQ SEQUENCE 185 AA; 21418 MW; B0BC7795 CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 185;  
Best Local Similarity 100.0%; Pred.No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 71 INKDDV 76  
Qy 16 INKDDV 21  
RESULT 3  
ID YXJG\_BACSU STANDARD; PRT; 334 AA.  
AC P42318;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 38.0 KD PROTEIN IN KATB 3' REGION.  
GN YXJG OR N15NR.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RA YOSHIDA K.-I., SANO H., SEKI S., ODA M., FUJIMURA M., FUJITA Y.;  
RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS YXJH.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D83026; D1012375;  
DR EMBL: Z99123; E1186395;  
DR SUBTILIS: BG11156; YXJG.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 334 AA; 38017 MW; CF69B60B CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 334;  
Best Local Similarity 100.0%; Pred.No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
calcium-binding proteins.";  
EUR. J. BIOCHEM. 192:387-399(1990).  
-!- FUNCTION: LIKE PARVALBUMINS, SCP'S SEEM TO BE MORE ABUNDANT IN  
CC FAST CONTRACTING MUSCLES, BUT NO FUNCTIONAL RELATIONSHIP CAN BE  
CC ESTABLISHED FROM THIS DISTRIBUTION.  
CC -!- THE SARCOPLASMIC CALCIUM-BINDING PROTEINS ARE ABUNDANT IN THE  
CC MUSCLE OF ARTHROPODS, MOLLUSCS, ANNELIDS, AND PROTOCHORDATES.  
CC -!- THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING SITES;  
CC POTENTIAL SITE 4 HAS LOST AFFINITY FOR CALCIUM.  
CC -!- THERE ARE 7 DIFFERENT SCP'S IN AMPHIOXUS. THE SEQUENCE SHOWN  
CC HERE IS THAT OF SCP I.  
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
DR PIR: B24479; B24479.  
DR PIR: S13182; S13182.  
DR PIR: S13184; S13184.  
DR PIR: S13185; S13185.  
DR PROSITE: PS00018; EF\_HAND; 3.  
DR PFAM: PF00036; efhand; 1.  
DR HSP: P04570; 2SAS.  
KW MUSCLE PROTEIN; CALCIUM-BINDING; DUPLICATION.  
FT CA\_BIND 19 30 SITE 1.  
FT CA\_BIND 70 81 SITE 2.  
FT CA\_BIND 115 126 SITE 3.  
FT DOMAIN 150 161 ANCESTRAL CALCIUM SITE 4.  
FT VARIANT 20 20 Y -> M (IN SCP III).  
FT VARIANT 23 23 D -> N (IN SCP IV).  
SQ SEQUENCE 185 AA; 21418 MW; B0BC7795 CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 185;  
Best Local Similarity 100.0%; Pred.No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 71 INKDDV 76  
Qy 16 INKDDV 21  
RESULT 3  
ID YXJG\_BACSU STANDARD; PRT; 334 AA.  
AC P42318;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 38.0 KD PROTEIN IN KATB 3' REGION.  
GN YXJG OR N15NR.  
OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RA YOSHIDA K.-I., SANO H., SEKI S., ODA M., FUJIMURA M., FUJITA Y.;  
RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS YXJH.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D83026; D1012375;  
DR EMBL: Z99123; E1186395;  
DR SUBTILIS: BG11156; YXJG.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 334 AA; 38017 MW; CF69B60B CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 334;  
Best Local Similarity 100.0%; Pred.No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 54 IENDEI 59  
Qy 2 IENDEI 7  
RESULT 4  
ID K2C8\_BOVIN STANDARD; PRT; 370 AA.  
AC P05786;  
DT 01-NOV-1988 (REL. 09, CREATED)  
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE KERATIN, TYPE II CYTOSKELETAL 8 (CYTOKERATIN 8) (CYTOKERATIN A)  
DE (FRAGMENT).  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 86193259.  
RA MAGIN T.M., JORCANO J.L., FRANK W.W.;  
RT "Cytokeratin expression in simple epithelia. II. cDNA cloning and  
RT sequence characteristics of bovine cytokeratin A (no. 8).";  
RL DIFFERENTIATION 30:254-264(1986).  
CC -!- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.  
CC KERATIN 8 ASSOCIATES WITH KERATIN 18.  
CC -!- THERE ARE TWO TYPES OF CYTOSKELETAL AND MICROFIBRILLAR KERATIN,  
CC I (ACIDIC) AND II (NEUTRAL TO BASIC) (40-55 AND 56-70 KILODALTONS,  
CC RESPECTIVELY).  
CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: K03532; G163245;  
DR EMBL: X12877; G481;  
DR PIR: A25004; A25004.  
DR PROSITE: PS00236; IF; 1.  
DR PFAM: PF00038; filament; 1.  
KW INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; KERATIN.  
FT NON\_TER 1 1  
FT DOMAIN <1 297 ROD.  
FT DOMAIN 298 370 TAIL.  
FT DOMAIN <1 25 COIL 1A.  
FT DOMAIN 26 42 LINKER 1.  
FT DOMAIN 43 134 COIL 1B.  
FT DOMAIN 135 158 LINKER 12.  
FT DOMAIN 159 297 COIL 2.  
FT SITE 241 241 STUTTER.  
SQ SEQUENCE 370 AA; 42395 MW; 6D2CD845 CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 370;  
Best Local Similarity 100.0%; Pred.No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 340 FSRSS 345  
Qy 28 FSRSS 33  
RESULT 5  
ID K2C8\_HUMAN STANDARD; PRT; 482 AA.  
AC P05787; Q14716; Q14717; Q14099;  
DT 01-NOV-1988 (REL. 09, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE KERATIN, TYPE II CYTOSKELETAL 8 (CYTOKERATIN 8) (K8) (CK 8).  
GN KRT8 OR CYK8.  
OS HOMO SAPIENS (HUMAN).



OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 90258929.  
RA YAMAMOTO R., KAO L.C., MCKNIGHT C.E., STRAUSS J.F. III;  
RT "Cloning and sequence of cDNA for human placental cytokeratin 8.  
RT Regulation of the mRNA in trophoblastic cells by cAMP."  
RL MOL. ENDOCRINOL. 4:370-374(1990).  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 90215304.  
RA KRAUSS S., FRANKE W.W.;  
RT "Organization and sequence of the human gene encoding cytokeratin 8."  
RL GENE 86:241-249(1990).  
RN [3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE: 91145351.  
RA WASEEM A., ALEXANDER C.M., STEEL J.B., LANE E.B.;  
RT "Embryonic simple epithelial keratins 8 and 18: chromosomal location  
RT emphasizes difference from other keratin pairs."  
RL NEW BIOL. 2:464-478(1990).  
RN [4]  
RN SEQUENCE FROM N.A., AND PHOSPHORYLATION.  
RX MEDLINE: 97207326.  
RA KU N.-O., OMARY M.B.;  
RT "Phosphorylation of human keratin 8 in vivo at conserved head domain  
RT serine 23 and at epidermal growth factor-stimulated tail domain  
RT serine 431."  
RL J. BIOL. CHEM. 272:7556-7564(1997).  
RN [5]  
RN SEQUENCE OF 1-230 FROM N.A.  
RX MEDLINE: 89261783.  
RA KULESH D.A., CECENA G., DARMON Y.M., VASSEUR M., OSHIMA R.G.;  
RT "Posttranslational regulation of keratins: degradation of mouse and  
RT human keratins 18 and 8."  
RL MOL. CELL. BIOL. 9:1553-1565(1989).  
RN [6]  
RN SEQUENCE OF 204-482 FROM N.A.  
RX MEDLINE: 87134779.  
RA LEUBE R.E., BOSCH F.X., ROMANO V., ZIMBELMANN R., HOFER H.,  
RA FRANKE W.W.;  
RT "Cytokeratin expression in simple epithelia. III. Detection of mRNAs  
RT encoding human cytokeratins nos. 8 and 18 in normal and tumor cells  
RT by hybridization with cDNA sequences in vitro and in situ."  
RL DIFFERENTIATION 33:69-85(1986).  
RN [7]  
RN PARTIAL SEQUENCE.  
RC TISSUE-COLON CARCINOMA;  
RX MEDLINE: 97295306.  
RA JI H., REID G.E., MORITZ R.L., EDDIS J.S., BURGESS A.W., SIMPSON R.J.;  
RT "A two-dimensional gel database of human colon carcinoma proteins."  
RL ELECTROPHORESIS 18:603-613(1997).  
CC KERATIN 8 ASSOCIATES WITH KERATIN 18.  
CC -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.  
CC -1- PTM: PHOSPHORYLATION ON SERINE RESIDUES ENHANCED DURING EGF  
CC STIMULATION AND MITOSIS.  
CC -1- THERE ARE TWO TYPES OF CYTOSKELETAL AND MICROFIBRILLAR KERATIN:  
CC I (ACIDIC; 40-55 KD) [K9 TO K20] AND II (NEUTRAL TO BASIC; 56-70  
CC KD) [K1 TO K8].  
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M34225; G181400; -.  
DR EMBL: M34482; G181573; -.  
DR EMBL: X74929; G400416; -.

DR EMBL: U76549; G1673575; -.  
DR EMBL: M26324; G307083; -.  
DR EMBL: X98614; E251594; -.  
DR EMBL: M26512; G553163; -.  
DR EMBL: X12882; G30313; -.  
DR EMBL: X74981; G400418; -.  
DR PIR: JS0487; JS0487.  
DR PIR: A34720; A34720.  
DR PIR: S06888; S06888.  
DR MIM: 148060; -.  
DR SWISS-2DPAGE; P05787; HUMAN.  
DR PROSITE; PS00226; IF; 1.  
DR PFAM; PF00038; filament; 1.  
KW INTERMEDIATE FILAMENT; COILED COIL; HEPTAD REPEAT PATTERN; KERATIN;  
KW PHOSPHORYLATION.  
FT INIT\_MET 0  
FT DOMAIN 1 89 HEAD.  
FT DOMAIN 90 397 ROD.  
FT DOMAIN 398 482 TAIL.  
FT DOMAIN 90 125 COIL 1A.  
FT DOMAIN 126 142 LINKER 1.  
FT DOMAIN 143 234 COIL 1B.  
FT DOMAIN 235 258 LINKER 12.  
FT DOMAIN 259 397 COIL 2.  
FT SITE 341 341 STUTTER.  
FT MOD\_RES 23 23 PHOSPHORYLATION.  
FT MOD\_RES 431 431 PHOSPHORYLATION (BY MAPK AND CAM-KINASE II).  
FT CONFLICT 76 76 V -> S (IN REF. 1).  
FT CONFLICT 200 200 D -> DVD (IN G400418).  
FT CONFLICT 231 231 I -> L (IN E251594).  
FT CONFLICT 256 256 D -> E (IN REF. 1).  
FT CONFLICT 383 383 L -> M (IN E251594).  
FT CONFLICT 386 386 E -> D (IN REF. 1).  
FT CONFLICT 400 400 R -> P (IN REF. 1).  
FT CONFLICT 416 416 G -> S (IN REF. 3 AND 4).  
FT CONFLICT 428 428 G -> D (IN REF. 1 AND 6).  
FT CONFLICT 429 432 LTSP -> SOA (IN REF. 2).  
FT CONFLICT 430 430 T -> A (IN E251594).  
FT CONFLICT 431 431 S -> D (IN REF. 1 AND 6).  
SQ SEQUENCE 482 AA; 53543 MW; 07D9E1C1 CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 482;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 451 FSRSTS 456  
| | | | |

QY 28 FSRSTS 33

RESULT 6 STANDARD; PRT; 670 AA.  
ID SR72\_CANFA  
AC P33731;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE SIGNAL RECOGNITION PARTICLE 72 KD PROTEIN (SRP72).  
GN SRP72.  
OS CANIS FAMILIARIS (DOG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC CARNIVORA; FISSIPEDIA; CANIDAE; CANIS.  
RN [1]  
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE: 93273803.  
RA LUETCKE H., PREHN S., ASHFORD A.J., REMUS M., FRANK R.,  
RA DOBERSTEIN B.;  
RT "Assembly of the 68- and 72-kD proteins of signal recognition  
RT particle with 7S RNA."  
RL J. CELL BIOL. 121:977-985(1993).  
CC -1- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY HAS A CRUCIAL ROLE  
CC IN TARGETING SECRETORY PROTEINS TO THE ROUGH ENDOPLASMIC  
CC RETICULUM MEMBRANE. SRP72 BINDS THE 7S RNA ONLY IN PRESENCE OF



CC SRP68. THIS RIBONUCLEOPROTEIN COMPLEX MIGHT INTERACT DIRECTLY WITH  
CC THE DOCKING PROTEIN IN THE ER MEMBRANE AND POSSIBLY PARTICIPATE  
CC IN THE ELONGATION ARREST FUNCTION.  
CC -!- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA MOLECULE  
CC OF 300 NUCLEOTIDES AND SIX PROTEIN SUBUNITS: SRP72, SRP68, SRP54,  
CC SRP19, SRP14 AND SRP9.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- DOMAIN: THE C-TERMINAL IS ESSENTIAL FOR THE INTERACTION WITH THE  
CC SRP68/7S RNA COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE SRP72 FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X67813; G297768; -;  
DR PIR; S32167; S32167;  
DR PIR; A40692; A40692;  
KW SIGNAL RECOGNITION PARTICLE; RIBONUCLEOPROTEIN.  
FT INIT\_MET 0 0 BLOCKED.  
FT MOD\_RES 1 1  
FT DOMAIN 551 560 POLY-LYS.  
FT DOMAIN 661 664 POLY-LYS.  
SQ SEQUENCE 670 AA; 74362 MW; 9CB38962 CRC32;  
  
Query Match 17.1%; Score 6; DB 1; Length 670;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 39 INKDDV 44  
| | | | |  
QY 16 INKDDV 21  
  
RESULT 7  
ID SR72\_HUMAN STANDARD; PRT; 670 AA.  
AC O76094;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DE 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE SIGNAL RECOGNITION PARTICLE 72 KD PROTEIN (SRP72).  
GN SRP72.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GONDA K., ZWIEB C.;  
RT "Protein SRP72 sequence of human signal recognition particle.";  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA UTZ P.J., HOTTELET M., MILLER I.J., ANDERSON P.;  
RT "Sequence of human signal recognition particle (SRP) 72.";  
RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY HAS A CRUCIAL ROLE  
CC IN TARGETING SECRETORY PROTEINS TO THE ROUGH ENDOPLASMIC  
CC RETICULUM MEMBRANE. SRP72 BINDS THE 7S RNA ONLY IN PRESENCE OF  
CC SRP68. THIS RIBONUCLEOPROTEIN COMPLEX MIGHT INTERACT DIRECTLY WITH  
CC THE DOCKING PROTEIN IN THE ER MEMBRANE AND POSSIBLY PARTICIPATE  
CC IN THE ELONGATION ARREST FUNCTION.  
CC -!- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA MOLECULE  
CC OF 300 NUCLEOTIDES AND SIX PROTEIN SUBUNITS: SRP72, SRP68, SRP54,  
CC SRP19, SRP14 AND SRP9.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- DOMAIN: THE C-TERMINAL IS ESSENTIAL FOR THE INTERACTION WITH THE  
CC SRP68/7S RNA COMPLEX (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SRP72 FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF077019; G3335650; -;  
DR EMBL; AF069765; G3243033; -;  
DR MIM; 602122; -;  
KW SIGNAL RECOGNITION PARTICLE; RIBONUCLEOPROTEIN.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT DOMAIN 551 560 POLY-LYS.  
FT DOMAIN 661 664 POLY-LYS.  
SQ SEQUENCE 670 AA; 74475 MW; 2EEF8AB8 CRC32;  
  
Query Match 17.1%; Score 6; DB 1; Length 670;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 39 INKDDV 44  
| | | | |  
QY 16 INKDDV 21  
  
RESULT 8  
ID MGR3\_HUMAN STANDARD; PRT; 877 AA.  
AC Q14832;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE METABOTROPIC GLUTAMATE RECEPTOR 3 PRECURSOR.  
GN GRM3 OR MGLUR3.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MAKOFF A., VOLPE F., LELCHUK R., HARRINGTON K., EMSON P.;  
RT "Molecular characterization and localization of human metabotropic  
RT glutamate receptor type 3.";  
RL BRAIN RES. MOL. BRAIN RES. 40:55-63(1996).  
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR  
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE  
CC ACTIVITY.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.  
CC STRONGEST, TO MGLUR2.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X77748; E99039; -;  
DR GCRDB; GCR\_2070; -;  
DR MIM; 601115; -;  
DR PROSITE; PS00979; G\_PROTEIN\_RECEP\_F3\_1; 1.  
DR PROSITE; PS00980; G\_PROTEIN\_RECEP\_F3\_2; 1.  
DR PROSITE; PS00981; G\_PROTEIN\_RECEP\_F3\_3; 1.  
DR PFAM; PF00003; 7tm\_3; 1.  
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL;  
KW MULTIGENE FAMILY.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 877 METABOTROPIC GLUTAMATE RECEPTOR 3.  
FT DOMAIN 21 574 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 575 597 I (POTENTIAL).  
FT -----

FT DOMAIN 598 611 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 612 632 II (POTENTIAL).  
FT DOMAIN 633 643 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 644 662 III (POTENTIAL).  
FT DOMAIN 663 686 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 687 707 IV (POTENTIAL).  
FT DOMAIN 708 732 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 733 754 V (POTENTIAL).  
FT DOMAIN 755 767 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 768 790 VI (POTENTIAL).  
FT DOMAIN 791 800 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 801 826 VII (POTENTIAL).  
FT DOMAIN 827 877 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 207 207 POTENTIAL.  
FT CARBOHYD 290 290 POTENTIAL.  
FT CARBOHYD 412 412 POTENTIAL.  
FT CARBOHYD 437 437 POTENTIAL.  
SQ SEQUENCE 877 AA; 98619 MW; B5DOAE5 CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 877;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 76 EINKDD 81 J  
QY 15 EINKDD 20  
RESULT 9  
ID ECLC-BOVIN STANDARD; PRT; 903 AA.  
AC P54281;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE CHANNEL).  
DE CHANNEL).  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=TRACHEA;  
RC MEDLINE; 96125078.  
RA CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,  
RA ARRATTE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;  
RT "Cloning of an epithelial chloride channel from bovine trachea.";  
RL J. BIOL. CHEM. 270:31016-31026(1995).  
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
CC TRANSEPITHELIAL TRANSPORT.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- TISSUE SPECIFICITY: TRACHEA.  
CC -!- PTM: PHOSPHORYLATED BY CAM-KINASE II.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U36445; G1184066;  
KW IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;  
KW CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.  
FT TRANSMEM 7 27 POTENTIAL.  
FT TRANSMEM 331 351 POTENTIAL.  
FT TRANSMEM 617 637 POTENTIAL.  
FT TRANSMEM 883 903 POTENTIAL.  
FT CARBOHYD 75 75 POTENTIAL.  
FT CARBOHYD 278 278 POTENTIAL.  
FT CARBOHYD 360 360 POTENTIAL.

FT CARBOHYD 372 372 POTENTIAL.  
FT CARBOHYD 504 504 POTENTIAL.  
FT CARBOHYD 515 515 POTENTIAL.  
FT CARBOHYD 688 688 POTENTIAL.  
FT CARBOHYD 811 811 POTENTIAL.  
FT CARBOHYD 816 816 POTENTIAL.  
FT CARBOHYD 842 842 POTENTIAL.  
FT CARBOHYD 857 857 POTENTIAL.  
SQ SEQUENCE 903 AA; 100305 MW; 65E6EC1C CRC32;  
Query Match 17.1%; Score 6; DB 1; Length 903;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 717 NPPRPE 722  
QY 10 NPPRPE 15  
RESULT 10  
ID HMDH\_DROME STANDARD; PRT; 916 AA.  
AC P14773;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE 3-HYDROXY-3-METHYLGLUTARYL-COENZYME A REDUCTASE (EC 1.1.1.34) (HMG-COA REDUCTASE).  
DE REDUCTASE).  
GN HMGCR OR HMG-COAR.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88302188.  
RA GERTLER F.B., CHIU C.-Y., RICHTER-MANN L., CHIN D.J.;  
RT "Developmental and metabolic regulation of the Drosophila  
RT melanogaster 3-hydroxy-3-methylglutaryl coenzyme A reductase.";  
RL MOL. CELL. BIOL. 8:2713-2721(1988).  
CC -!- FUNCTION: SYNTHESIS OF MEVALONATE FOR THE PRODUCTION OF NON-  
CC STEROL ISOPRENOIDS, WHICH ARE ESSENTIAL FOR GROWTH  
CC DIFFERENTIATION.  
CC -!- CATALYTIC ACTIVITY: (R)-MEVALONATE + COA + 2 NADP(+) = (S)-3-  
CC HYDROXY-3-METHYLGLUTARYL-COA + 2 NADPH.  
CC -!- ENZYME REGULATION: THE ACTIVITY OF HMG-COA-REDUCTASE IS  
CC SUPPRESSED BY EXOGENOUS MEVALONATE.  
CC -!- PATHWAY: ISOPRENOID BIOSYNTHESIS.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC  
CC RETICULUM.  
CC -!- SIMILARITY: BELONGS TO THE HMG-COA REDUCTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M21329; G157624; -.  
DR FIR; S32572; S32572.  
DR FLYBASE; FBgn0001205; Hmgcr.  
DR PROSITE; PS00066; HMG\_COA\_REDUCTASE\_1; 1.  
DR PROSITE; PS00318; HMG\_COA\_REDUCTASE\_2; 1.  
DR PROSITE; PS01192; HMG\_COA\_REDUCTASE\_3; 1.  
DR PROSITE; PS00655; HMG\_COA\_REDUCTASE\_4; 1.  
DR PFAM; PF00368; HMG-CoA\_red1; 1.  
DR PFAM; PF00369; HMG-CoA\_red2; 1.  
DR HSP; P19656; IAPH.  
DR OXIDOREDUCTASE; GLYCOPROTEIN; ENDOPLASMIC RETICULUM; TRANSMEMBRANE;  
KW ISOPRENE BIOSYNTHESIS; NADP.  
FT DOMAIN 1 380 MEMBRANE-BOUND.  
FT DOMAIN 381 494 LINKER.

FT DOMAIN 495 825 CATALYTIC.  
FT TRANSMEM 11 33 POTENTIAL.  
FT TRANSMEM 100 127 POTENTIAL.  
FT TRANSMEM 133 159 POTENTIAL.  
FT TRANSMEM 167 191 POTENTIAL.  
FT TRANSMEM 202 226 POTENTIAL.  
FT TRANSMEM 234 262 POTENTIAL.  
FT TRANSMEM 358 381 POTENTIAL.  
FT ACT\_SITE 582 582 BY SIMILARITY.  
FT ACT\_SITE 789 789 BY SIMILARITY.  
FT ACT\_SITE 888 888 GENERAL BASE (BY SIMILARITY).  
FT CARBOHYD 338 338 POTENTIAL.  
FT CARBOHYD 342 342 POTENTIAL.  
FT CARBOHYD 439 439 POTENTIAL.  
FT CARBOHYD 471 471 POTENTIAL.  
FT CARBOHYD 793 793 POTENTIAL.  
FT CARBOHYD 798 798 POTENTIAL.  
FT CARBOHYD 892 892 POTENTIAL.  
FT CARBOHYD 906 906 POTENTIAL.  
SQ SEQUENCE 916 AA; 98295 MW; C11833A2 CRC32;  
  
Query Match 17.1%; Score 6; DB 1; Length 916;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 630 WIENDE 635  
|||||  
Qy 1 WIENDE 6  
  
RESULT 11  
ID CLH\_CAEEL STANDARD; PRT; 1681 AA.  
AC P34574;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PROBABLE CLATHRIN HEAVY CHAIN.  
GN T20G5.1.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SEGERNENTEA; RHABDITIA; RHABDITIDA;  
RN RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
[1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=BRISTOL NZ;  
RA BERKS M., SMITH A.;  
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYHEDRAL COAT OF  
CC COATED PITS AND VESICLES (BY SIMILARITY).  
CC -!- SUBUNIT: CLATHRIN TRISKELIONS, COMPOSED OF 3 HEAVY CHAINS AND 3  
CC LIGHT CHAINS, ARE THE BASIC SUBUNITS OF THE CLATHRIN COAT  
CC (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND  
CC VESICLES (BY SIMILARITY).  
CC -!- DOMAIN: THE C-TERMINAL THIRD OF THE HEAVY CHAINS FORMS THE HUB OF  
CC THE TRISKELION. THIS REGION CONTAINS THE TRIMERIZATION DOMAIN AND  
CC THE LIGHT-CHAIN BINDING DOMAIN INVOLVED IN THE ASSEMBLY OF THE  
CC CLATHRIN LATTICE.  
CC -!- SIMILARITY: BELONGS TO THE CLATHRIN HEAVY CHAIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z30423; G458481; -.  
DR PIR; S42369; S42369.  
DR WORMPEP; T20G5.1; CE00480.  
DR PFAM; PF00637; Clathrin\_repeat; 7.  
KW HYPOTHETICAL PROTEIN; COATED PITS.  
SQ SEQUENCE 1681 AA; 191541 MW; 6CD2B462 CRC32;

Query Match 17.1%; Score 6; DB 1; Length 1681;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1256 KOVCF5 1261  
|||||  
Qy 24 KOVCF5 29  
  
RESULT 12  
ID RPO\_PVNR STANDARD; PRT; 1967 AA.  
AC P17955;  
DT 01-NOV-1990 (REL. 16, CREATED)  
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE RNA REPLICATION PROTEIN (147 KD PROTEIN) (ORF 1) [CONTAINS: RNA-  
DE DIRECTED RNA POLYMERASE (EC 2.7.7.48); PROBABLE HELICASE].  
OS POTATO VIRUS M (STRAIN RUSSIAN) (PVM).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; CARLAVIRUS.  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 91116326.  
RX ZAVRIEV S.K., KANYUKA K.V., LEVAY K.E.;  
RT "The genome organization of potato virus M RNA."  
RL J. GEN. VIROL. 72:9-14(1991).  
CC -!- FUNCTION: RNA-REPLICATION. THE CENTRAL PART OF THIS PROTEIN  
CC POSSIBLY FUNCTIONS AS A ATP-BINDING HELICASE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X53062; -; NOT\_ANNOTATED\_CDS.  
DR PIR; S21601; S21601.  
DR PIR; PNO093; PNO093.  
DR HSP; P03069; 22TA.  
KW ATP-BINDING; HELICASE; RNA REPLICATION; RNA-DIRECTED RNA POLYMERASE;  
KW TRANSFERASE.  
FT NP\_BIND 1165 1172 ATP (POTENTIAL).  
SQ SEQUENCE 1967 AA; 223207 MW; CBD4EFCDC32;  
  
Query Match 17.1%; Score 6; DB 1; Length 1967;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 1324 EINKDD 1329  
|||||  
Qy 15 EINKDD 20  
  
RESULT 13  
ID POLN\_RUBYV STANDARD; PRT; 2205 AA.  
AC P13889;  
DT 01-JAN-1990 (REL. 13, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)  
DE NONSTRUCTURAL POLYPROTEIN [CONTAINS: NONSTRUCTURAL PROTEINS NSP1 TO  
DE NSP4].  
OS RUBELLA VIRUS (STRAIN THERIEN).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; TOGAVIRIDAE;  
OC RUBIVIRUS.  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE; 90281585.  
RX DOMINGUEZ G., WANG C.Y., FREY T.K.;  
RT "Sequence of the genome RNA of rubella virus: evidence for genetic  
RT rearrangement during togavirus evolution."  
RL VIROLOGY 177:225-258(1990).

RN [2] SEQUENCE OF 1737-2205 FROM N.A.  
RX MEDLINE: 88226020.  
RY FREY T.K., MARR L.D.:  
RA "Sequence of the region coding for virion proteins C and E2 and the  
RT carboxy terminus of the nonstructural proteins of rubella virus:  
RL comparison with alphaviruses.";  
RM GENE 62:85-99(1988).  
RN CC -1- PPM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M15240; G333972; -  
DR PTR: A35320; MNWVRN.  
KW POLYPEPTIDE; NONSTRUCTURAL PROTEIN.  
FT CHAIN 1 ? NONSTRUCTURAL PROTEIN NSP1.  
FT CHAIN ? ? NONSTRUCTURAL PROTEIN NSP2.  
FT CHAIN ? ? NONSTRUCTURAL PROTEIN NSP3.  
FT CHAIN ? 2205 NONSTRUCTURAL PROTEIN NSP4.  
SQ SEQUENCE 2205 AA; 240220 MW; E2158EB4 CRC32;  
  
Query Match 17.1%; Score 6; DB 1; Length 2205;  
Best Local Similarity 100.0%; Pred. No. 2.56e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 474 PPRPEI 479  
QY 11 PPRPEI 16  
|||||  
  
RESULT 14  
ID NOTC\_DROME STANDARD; PRT: 2703 AA.  
AC P07207; P04154;  
DT 01-NOV-1986 (REL. 03, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE NEUROGENIC LOCUS NOTCH PROTEIN PRECURSOR.  
GN N.  
OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;  
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;  
OC DROSOPHILIDAE; DROSOPHILA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 86079539.  
RA WHARTON K.A., JOHANSEN K.M., XU T., ARTAVANIS-TSAKONAS S.;  
RT "Nucleotide sequence from the neurogenic locus notch implies a gene  
RT product that shares homology with proteins containing EGF-like  
RT repeats.";  
RL CELL 43:567-581(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-OREGON-R;  
RX MEDLINE: 87064624.  
RA KIDD S., KELLEY M.R., YOUNG M.W.;  
RT "Sequence of the notch locus of Drosophila melanogaster; relationship  
RT of the encoded protein to mammalian clotting and growth factors.";  
RM MOL. CELL. BIOL. 6:3094-3108(1986).  
RN [3]  
RP SEQUENCE OF 2505-2611 FROM N.A.  
RX MEDLINE: 85099329.  
RA WHARTON K.A., YEDVOBNICK B., FINNERTY V.G., ARTAVANIS-TSAKONAS S.;  
RT "opa: a novel family of transcribed repeats shared by the Notch locus  
RT and other developmentally regulated loci in D. melanogaster.";  
RM CELL 40:55-62(1985).  
RN [4]  
RP SEQUENCE OF 1-8 FROM N.A.

RX MEDLINE: 87257846.  
RA KELLEY M.R., KIDD S., BERG R.L., YOUNG M.W.;  
RT "Restriction of P-element insertions at the Notch locus of Drosophila  
RT melanogaster.";  
RM MOL. CELL. BIOL. 7:1545-1548(1987).  
RN [5]  
RP REVIEW.  
RA HARRIS W.A.;  
RT "Many cell types specified by Notch function.";  
RM CURR. BIOL. 1:120-122(1991).  
RN CC -1- FUNCTION: NOTCH PROTEIN IS ESSENTIAL FOR PROPER DIFFERENTIATION OF  
CC ECTODERM.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO THE INNER PART  
CC OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS,  
CC THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.  
CC -1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.  
CC -1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
CC -1- SIMILARITY: CONTAINS 6 ANK REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M16152; G157988; -  
DR EMBL: M16153; G157988; JOINED.  
DR EMBL: M16149; G157988; JOINED.  
DR EMBL: M16150; G157988; JOINED.  
DR EMBL: M16151; G157988; JOINED.  
DR EMBL: K03508; G157993; -  
DR EMBL: M13689; G157993; JOINED.  
DR EMBL: K03507; G157993; JOINED.  
DR EMBL: M12175; G950317; -  
DR EMBL: M16025; G157995; -  
DR PIR: A24420; A24420.  
DR PIR: A24768; A24768.  
DR PIR: A05267; A05267.  
DR FLYBASE: FBgn0004647; N.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 22.  
DR PROSITE: PS00022; EGF\_1; 34.  
DR PROSITE: PS01186; EGF\_2; 28.  
DR PROSITE: PS01187; EGF\_CA; 22.  
DR PFAM: PF00008; EGF; 36.  
DR PFAM: PF00023; ank; 6.  
DR PFAM: PF00066; notch; 3.  
DR HSP: P00740; 11XA.  
KW DIFFERENTIATION; NEUROGENESIS; REPEAT; ANK REPEAT; EGF-LIKE DOMAIN;  
KW TRANSMEMBRANE; SIGNAL; GLYCOPROTEIN.  
FT SIGNAL 1 44 POTENTIAL.  
FT CHAIN 45 2703 NEUROGENIC LOCUS NOTCH PROTEIN.  
FT DOMAIN 45 1745 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1746 1766 POTENTIAL.  
FT DOMAIN 1767 2703 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 58 1451 36 X EGF-TYPE REPEATS.  
FT DOMAIN 58 95 EGF-LIKE 1.  
FT DOMAIN 96 136 EGF-LIKE 2.  
FT DOMAIN 139 176 EGF-LIKE 3.  
FT DOMAIN 177 215 EGF-LIKE 4.  
FT DOMAIN 217 253 EGF-LIKE 5.  
FT DOMAIN 255 291 EGF-LIKE 6.  
FT DOMAIN 293 329 EGF-LIKE 7.  
FT DOMAIN 331 370 EGF-LIKE 8.  
FT DOMAIN 372 408 EGF-LIKE 9.  
FT DOMAIN 409 447 EGF-LIKE 10.  
FT DOMAIN 449 486 EGF-LIKE 11.  
FT DOMAIN 488 524 EGF-LIKE 12.  
FT DOMAIN 526 562 EGF-LIKE 13.  
FT DOMAIN 564 600 EGF-LIKE 14.  
FT DOMAIN 600 600 CALCIUM-BINDING (POTENTIAL).



**This Page Blank (uspto)**

\*\*\*\*\*  
WIREHOLE  
\*\*\*\*\*  
(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:06:25 1999; Maspar time 7.78 Seconds  
Tabular output not generated. 245.544 Million cell updates/sec

Title: >US-09-049-696-48  
Description: (1-35) from US09049696.pep  
Perfect Score: 35  
Sequence: 1 WIENDEIQWNPPRPEINKDDVQHKVCFRTSSGG 35

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: sptrembl9  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 2.705; Variance 0.430; scale 6.293

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	9	25.7	913 11	GOB-5 PROTEIN.	3.45e-05
2	7	20.0	287 1	METHYLVILOGEN-REDUCIN	2.01e-01
3	6	17.1	84 5	COSMID R02E12.	1.00e+01
4	6	17.1	131 10	ESRIG2 (ESR1C1) PROTEIN	1.00e+01
5	6	17.1	170 2	PUTATIVE MINOR FIMBRIA	1.00e+01
6	6	17.1	184 1	184RA LONG HYPOTHETICA	1.00e+01
7	6	17.1	218 4	PUTATIVE PEROXISOME MI	1.00e+01
8	6	17.1	249 5	SIMILARITY TO TOXOCARA	1.00e+01
9	6	17.1	263 3	SCEB PRECURSOR.	1.00e+01
10	6	17.1	293 13	HYPOTHETICAL 32.5 K D P	1.00e+01
11	6	17.1	314 5	K01D12-1 PROTEIN.	1.00e+01
12	6	17.1	376 5	COSMID F13B9.	1.00e+01
13	6	17.1	378 3	SAGA.	1.00e+01
14	6	17.1	393 5	E01G4.1 PROTEIN.	1.00e+01
15	6	17.1	423 1	HOMOSERINE DEHYDROGENA	1.00e+01
16	6	17.1	440 1	440AA LONG HYPOTHETICA	1.00e+01
17	6	17.1	460 5	F32A7.3B PROTEIN.	1.00e+01
18	6	17.1	469 9	ORF29.	1.00e+01
19	6	17.1	500 4	LIPRIN-ALPHA4 (FRAGMEN	1.00e+01
20	6	17.1	505 8	MATURASE.	1.00e+01

21	6	17.1	505 1	Q28927	1.00e+01
22	6	17.1	574 1	Q58327	1.00e+01
23	6	17.1	603 10	Q22787	1.00e+01
24	6	17.1	771 5	Q27338	1.00e+01
25	6	17.1	776 5	Q09501	1.00e+01
26	6	17.1	794 6	Q18742	1.00e+01
27	6	17.1	820 6	Q18743	1.00e+01
28	6	17.1	905 6	Q18741	1.00e+01
29	6	17.1	1451 5	Q61779	1.00e+01
30	6	17.1	1680 10	Q64588	1.00e+01
31	6	17.1	1968 14	Q89548	1.00e+01
32	6	17.1	2115 14	Q86500	1.00e+01
33	6	17.1	2116 14	Q40955	1.00e+01
34	6	17.1	2531 5	Q60004	1.00e+01
35	6	17.1	4377 4	Q12955	1.00e+01
36	5	14.3	20 14	Q65539	3.35e+02
37	5	14.3	26 10	Q38753	3.35e+02
38	5	14.3	26 10	Q38747	3.35e+02
39	5	14.3	43 11	Q88288	3.35e+02
40	5	14.3	43 11	Q88289	3.35e+02
41	5	14.3	53 14	Q86960	3.35e+02
42	5	14.3	76 14	Q68828	3.35e+02
43	5	14.3	82 14	Q76706	3.35e+02
44	5	14.3	82 14	Q71006	3.35e+02
45	5	14.3	85 14	Q72968	3.35e+02
46	5	14.3	90 14	Q74932	3.35e+02
47	5	14.3	91 14	Q68936	3.35e+02
48	5	14.3	92 5	Q27441	3.35e+02
49	5	14.3	93 14	Q70554	3.35e+02
50	5	14.3	93 14	Q77563	3.35e+02
51	5	14.3	93 14	Q67603	3.35e+02
52	5	14.3	95 14	Q77320	3.35e+02
53	5	14.3	95 14	Q77323	3.35e+02
54	5	14.3	98 14	Q77565	3.35e+02
55	5	14.3	102 14	Q90283	3.35e+02
56	5	14.3	109 2	Q53698	3.35e+02
57	5	14.3	110 9	Q80048	3.35e+02
58	5	14.3	111 2	Q89648	3.35e+02
59	5	14.3	111 14	Q76226	3.35e+02
60	5	14.3	115 1	Q57851	3.35e+02
61	5	14.3	115 14	Q39126	3.35e+02
62	5	14.3	117 5	Q94257	3.35e+02
63	5	14.3	119 6	Q95315	3.35e+02
64	5	14.3	119 5	Q17587	3.35e+02
65	5	14.3	122 14	Q86131	3.35e+02
66	5	14.3	122 9	Q38286	3.35e+02
67	5	14.3	130 1	Q59275	3.35e+02
68	5	14.3	131 6	Q92261	3.35e+02
69	5	14.3	134 2	Q53960	3.35e+02
70	5	14.3	135 1	Q73973	3.35e+02
71	5	14.3	140 1	Q58657	3.35e+02
72	5	14.3	140 14	Q68257	3.35e+02
73	5	14.3	140 2	Q57775	3.35e+02
74	5	14.3	140 14	Q88258	3.35e+02
75	5	14.3	140 14	Q86252	3.35e+02
76	5	14.3	141 2	Q54273	3.35e+02
77	5	14.3	145 2	Q05984	3.35e+02
78	5	14.3	146 5	Q09935	3.35e+02
79	5	14.3	147 1	Q50252	3.35e+02
80	5	14.3	148 2	Q93308	3.35e+02
81	5	14.3	149 5	Q52228	3.35e+02
82	5	14.3	149 5	Q52527	3.35e+02
83	5	14.3	151 2	Q78186	3.35e+02
84	5	14.3	155 2	Q50003	3.35e+02
85	5	14.3	157 5	Q26067	3.35e+02
86	5	14.3	157 2	Q47991	3.35e+02
87	5	14.3	159 10	Q41828	3.35e+02
88	5	14.3	159 5	Q18089	3.35e+02
89	5	14.3	162 5	Q17058	3.35e+02
90	5	14.3	163 3	Q13291	3.35e+02
91	5	14.3	167 10	Q41893	3.35e+02
92	5	14.3	168 10	Q41892	3.35e+02
93	5	14.3	172 2	Q50191	3.35e+02

THYMIDINE PHOSPHORYLAS  
HYPOTHETICAL PROTEIN M  
F4P9.13 PROTEIN.  
KSR-1.  
HYPOTHETICAL 89.0 KD P  
LU-ECAM-1.  
LU-ECAM-1.  
LU-ECAM-1.  
K10E9.1 PROTEIN.  
T2F13.11 PROTEIN.  
ORF1 223K PROTEIN.  
M33 RNA FOR A NONSTRUC  
NONSTRUCTURAL PROTEIN.  
NOTCH HOMOLOG.  
ANKYRIN G.  
(EBV) GENOME. THE COMP  
ASPI8 MADS BOX (FRAGMEN  
GLUTAMINE FRUCTOSE-6-P  
GLUTAMINE FRUCTOSE-6-P  
E1 (FRAGMENT).  
DNA FOR MTRIII REGION.  
ENVELOPE PROTEIN (FRAG  
ENVELOPE GLYCOPROTEIN  
ENVELOPE GLYCOPROTEIN  
ENVELOPE GLYCOPROTEIN  
RNA POLYMERASE (FRAGME  
NEUROPEPTIDE Y PRECURS  
ENVELOPE GLYCOPROTEIN  
ENVELOPE PROTEIN (FRAG  
REPLICATION-ASSOCIATED  
ENVELOPE PROTEIN (FRAG  
ENVELOPE PROTEIN (FRAG  
ENVELOPE PROTEIN (FRAG  
TAT PROTEIN.  
BETA-LACTAMASE (EC 3.5  
ORF 9.  
APT2 PROTEIN PRECURSOR  
ENVELOPE GLYCOPROTEIN  
115AA LONG HYPOTHETICA  
ENVELOPE GLYCOPROTEIN  
COSMID K04A8.  
ERYTHROID ISOFORM PROT  
C15H11.8 PROTEIN.  
G PROTEIN (FRAGMENT).  
E11.  
130AA LONG HYPOTHETICA  
VILLIN (FRAGMENT).  
THIENAMYCIN CYCLASE.  
135AA LONG HYPOTHETICA  
140AA LONG HYPOTHETICA  
RNA POLYMERASE (FRAGME  
CONSERVED HYPOTHETICAL  
RNA POLYMERASE (FRAGME  
RNA POLYMERASE (FRAGME  
ORF141 PROTEIN.  
HYPOTHETICAL 17.2 KD P  
HYPOTHETICAL 17.1 KD P  
FORMATE HYDROGENLYASE  
POLYGLUTAMATE SYN  
SERINE PROTEINASE (FRA  
SERINE PROTEINASE (FRA  
NADH DEHYDROGENASE I C  
U1764T.  
MYOSIN REGULATORY LIGH  
CRT GENES.  
MADS BOX PROTEIN (FRAG  
C18E9.9 PROTEIN.  
C24A1.1 PROTEIN.  
CHITIN SYNTHASE 1.  
ZEM2 (FRAGMENT).  
PLASMID PFO11 REGULATO









532 5 14.3 781 2 Q52753 XYNB PRECURSOR. 3.35e+02 605 5 14.3 1008 5 077391 MALP6.4 PROTEIN. 3.35e+02  
533 5 14.3 782 5 Q61073 SARCOPLASMIC/ENDOPLASM 3.35e+02 606 5 14.3 1024 5 Q09570 HYPOTHETICAL 114.8 KD 3.35e+02  
534 5 14.3 786 5 Q76133 DSRC29A TYPE 2 PROTEIN 3.35e+02 607 5 14.3 1029 4 Q60900 SARCO/ENDOPLASMIC RETI 3.35e+02  
535 5 14.3 789 2 Q45134 ARABINOSIDASE. 3.35e+02 608 5 14.3 1032 4 Q13866 CR2 RECEPTOR PRECURSOR 3.35e+02  
536 5 14.3 789 5 Q16781 SIMILARITY TO MULTIPLE 3.35e+02 609 5 14.3 1036 5 Q76974 SARCO/ENDOPLASMIC RETI 3.35e+02  
537 5 14.3 792 5 Q23115 WOLC9.3 PROTEIN. 3.35e+02 610 5 14.3 1036 10 Q49323 PUTATIVE CELLULOSE SYN 3.35e+02  
538 5 14.3 792 10 Q22881 PUTATIVE POTASSIUM TRA 3.35e+02 611 5 14.3 1038 11 Q64517 SARCOENDOPLASMIC RETIC 3.35e+02  
539 5 14.3 793 5 Q18215 Y53C12B.1 PROTEIN. 3.35e+02 612 5 14.3 1045 4 Q76037 MUCIN PRECURSOR (FRAGM 3.35e+02  
540 5 14.3 794 10 Q22398 PUTATIVE POTASSIUM TRA 3.35e+02 613 5 14.3 1046 5 P91176 CODED FOR BY C. ELEGAN 3.35e+02  
541 5 14.3 795 5 Q42706 SERINE-THREONINE KINAS 3.35e+02 614 5 14.3 1047 5 Q61136 CALCIUM MOTIVE P-TYPE 3.35e+02  
542 5 14.3 796 9 Q38354 MINOR STRUCTURAL PROTE 3.35e+02 615 5 14.3 1048 3 P78981 NA+-ATPASE. 3.35e+02  
543 5 14.3 804 10 Q49619 HYPOTHETICAL 90.8 KD P 3.35e+02 616 5 14.3 1055 3 Q13397 P-TYPE ATPASE 1. 3.35e+02  
544 5 14.3 805 5 Q27412 SERINE THREONINE KINAS 3.35e+02 617 5 14.3 1066 5 Q27802 DYNEIN HEAVY CHAIN ISO 3.35e+02  
545 5 14.3 807 14 Q98330 M83 PROTEIN. 3.35e+02 618 5 14.3 1081 10 Q48946 CELLULOSE SYNTHASE CAT 3.35e+02  
546 5 14.3 808 2 Q70021 BETA-GLUCOSIDASE. 3.35e+02 619 5 14.3 1082 3 Q13398 P-TYPE ATPASE 2. 3.35e+02  
547 5 14.3 809 14 Q71155 VIRAL POLYPEPTIDE VP1. 3.35e+02 620 5 14.3 1087 5 Q26156 V-SERA 5. 3.35e+02  
548 5 14.3 810 14 Q90053 DR035.28P. 3.35e+02 621 5 14.3 1087 4 Q14212 EPSTEIN-BARR VIRUS COM 3.35e+02  
549 5 14.3 810 3 Q03388 SIMILAR TO KLUVVEROMYC 3.35e+02 622 5 14.3 1094 5 Q61618 DYNAMIN ASSOCIATED PRO 3.35e+02  
550 5 14.3 811 14 Q84677 STRUCTURAL PROTEIN VP1 3.35e+02 623 5 14.3 1098 2 Q48152 ADHESIN (HIA). 3.35e+02  
551 5 14.3 811 14 Q90125 C2-HC TYPE ZINC FINGER 3.35e+02 624 5 14.3 1101 5 Q27235 DFURINI-CRR. 3.35e+02  
552 5 14.3 815 11 P70589 ENVELOPE POLYPEPTIDE. 3.35e+02 625 5 14.3 1102 2 Q52298 PLASMIN DMYSHG6000 VIRU 3.35e+02  
553 5 14.3 819 5 P90882 CODED FOR BY C. ELEGAN 3.35e+02 626 5 14.3 1114 13 Q42408 DELTAETFI. 3.35e+02  
554 5 14.3 827 2 Q47871 XYLANASE. 3.35e+02 627 5 14.3 1119 5 Q19190 HYPOTHETICAL 128.9 KD 3.35e+02  
555 5 14.3 835 2 Q56889 INVASIN. 3.35e+02 628 5 14.3 1120 10 Q81509 T7M24.3 PROTEIN. 3.35e+02  
556 5 14.3 842 3 Q00204 PUTATIVE CYSTEINE PROT 3.35e+02 629 5 14.3 1123 5 Q23851 HYPOTHETICAL 127.0 KD 3.35e+02  
557 5 14.3 844 10 Q81767 HYPOTHETICAL 94.5 KD P 3.35e+02 630 5 14.3 1142 3 P87253 PROTEIN KINASE C HOMOL 3.35e+02  
558 5 14.3 844 10 Q64472 BZIP-LIKE PROTEIN. 3.35e+02 631 5 14.3 1165 5 Q20650 CODED FOR BY C. ELEGAN 3.35e+02  
559 5 14.3 846 14 Q89292 ENVELOPE POLYPEPTIDE. 3.35e+02 632 5 14.3 1174 3 Q42632 PROTEIN KINASE C. 3.35e+02  
560 5 14.3 847 4 Q16584 PROTEIN KINASE. 3.35e+02 633 5 14.3 1176 14 Q98587 HYPOTHETICAL 119.0 KD 3.35e+02  
561 5 14.3 851 10 Q22260 KINESIN-LIKE PROTEIN, 3.35e+02 634 5 14.3 1182 11 Q08996 MYELIN TRANSCRIPTION F 3.35e+02  
562 5 14.3 874 14 Q89279 CAPSID PROTEIN PRECURS 3.35e+02 635 5 14.3 1183 11 Q89040 PHOSPHOLIPASE C BETA 2 3.35e+02  
563 5 14.3 875 10 Q42739 PYRUVATE, ORTHOPHOSPHAT 3.35e+02 636 5 14.3 1187 11 P70475 NEURAL ZINC FINGER TRA 3.35e+02  
564 5 14.3 876 14 Q79670 ENVELOPE PROTEIN GP120 3.35e+02 637 5 14.3 1188 11 P97500 ZINC FINGER PROTEIN PN 3.35e+02  
565 5 14.3 890 2 Q34431 YLOB PROTEIN. 3.35e+02 638 5 14.3 1188 10 Q65430 GLU-RICH PROTEIN. 3.35e+02  
566 5 14.3 891 5 Q61361 CHOL BETA. 3.35e+02 639 5 14.3 1189 11 P70392 RAS PROTEIN-SPECIFIC G 3.35e+02  
567 5 14.3 891 5 Q93763 F53C11.5 PROTEIN. 3.35e+02 640 5 14.3 1208 4 Q75162 KIAA0675 PROTEIN. 3.35e+02  
568 5 14.3 892 5 Q27436 FURINI. 3.35e+02 641 5 14.3 1211 13 Q91086 PHOSPHOLIPASE C BETA. 3.35e+02  
569 5 14.3 895 5 Q02810 CODED FOR BY C. ELEGAN 3.35e+02 642 5 14.3 1213 3 Q14340 HYPOTHETICAL 137.8 KD 3.35e+02  
570 5 14.3 897 13 Q13098 MIDDLE MOLECULAR WEIGH 3.35e+02 643 5 14.3 1224 3 Q13637 WEB1 PROTEIN. 3.35e+02  
571 5 14.3 898 4 Q14864 CHLORIDE CHANNEL PROTE 3.35e+02 644 5 14.3 1227 5 Q21038 RAS-GRF2 (FRAGMENT). 3.35e+02  
572 5 14.3 900 10 Q23454 HYPOTHETICAL 104.2 KD 3.35e+02 645 5 14.3 1237 4 Q14827 REOVIRUS DRV ORF. 3.35e+02  
573 5 14.3 901 11 Q88860 CHLORIDE CHANNEL CACC. 3.35e+02 646 5 14.3 1241 14 Q85444 ARYLPHORIN RECEPTOR. 3.35e+02  
574 5 14.3 904 5 Q19857 F28C1.3 PROTEIN. 3.35e+02 647 5 14.3 1253 5 Q23809 ARYLPHORIN RECEPTOR. 3.35e+02  
575 5 14.3 907 10 Q42738 PYRUVATE, ORTHOPHOSPHAT 3.35e+02 648 5 14.3 1253 5 P91601 FURIN 1, ISOFORM DFURI 3.35e+02  
576 5 14.3 913 13 Q13099 MIDDLE MOLECULAR WEIGH 3.35e+02 649 5 14.3 1269 5 Q05817 CHROMOSOME XII COSMID 3.35e+02  
577 5 14.3 919 5 Q45374 F17B5.2 PROTEIN. 3.35e+02 650 5 14.3 1272 13 Q90924 TYPE IIS RESTRICTION E 3.35e+02  
578 5 14.3 922 10 Q41058 STARCH BRANCHING ENZYM 3.35e+02 651 5 14.3 1274 3 Q06673 KIAA0551 PROTEIN (FRAG 3.35e+02  
579 5 14.3 925 10 Q48580 PUTATIVE CU2+-TRANSPOR 3.35e+02 652 5 14.3 1279 2 Q28046 COSMID W03F11. 3.35e+02  
580 5 14.3 926 3 Q13305 PROTEASE 1. 3.35e+02 653 5 14.3 1280 5 Q01779 KIAA0324 (FRAGMENT). 3.35e+02  
581 5 14.3 928 2 Q86395 NPP47 PROTEIN. 3.35e+02 654 5 14.3 1288 4 Q15038 BACTERIOPHAGE (PHIC31) 3.35e+02  
582 5 14.3 938 6 Q77658 CYCLIC NUCLEOTIDE-GATE 3.35e+02 655 5 14.3 1294 2 Q86682 BACTERIOPHAGE (PHIC31) 3.35e+02  
583 5 14.3 947 5 Q15696 ORNITHINE DECARBOXYLAS 3.35e+02 656 5 14.3 1294 2 Q53942 EG-115C2.10 PROTEIN. 3.35e+02  
584 5 14.3 948 6 Q77659 CYCLIC NUCLEOTIDE-GATE 3.35e+02 657 5 14.3 1300 5 Q77261 F2566.9 PROTEIN. 3.35e+02  
585 5 14.3 948 2 Q30483 PUTATIVE REGULATORY PR 3.35e+02 658 5 14.3 1326 5 Q16928 KIAA0551 PROTEIN (FRAG 3.35e+02  
586 5 14.3 950 5 Q24145 TYROSINE KINASE. 3.35e+02 659 5 14.3 1333 4 Q60298 PROTEIN TYROSINE KINAS 3.35e+02  
587 5 14.3 952 13 Q07494 TYROSINE KINASE CEK6 R 3.35e+02 660 5 14.3 1338 5 Q23927 RESTRICTIN PRECURSOR. 3.35e+02  
588 5 14.3 952 6 Q77660 CYCLIC NUCLEOTIDE-GATE 3.35e+02 661 5 14.3 1353 13 Q00546 KOPF6.6 PROTEIN. 3.35e+02  
589 5 14.3 962 4 Q60763 P115. 3.35e+02 662 5 14.3 1355 5 Q16732 CODED FOR BY C. ELEGAN 3.35e+02  
590 5 14.3 962 14 Q89443 HELICASE. 3.35e+02 663 5 14.3 1355 5 Q16779 NEUROFASCIN PRECURSOR. 3.35e+02  
591 5 14.3 962 5 Q61360 CHOL ALPHA. 3.35e+02 664 5 14.3 1369 13 Q42414 GLUCOSYLTRANSFERASE-SI 3.35e+02  
592 5 14.3 966 5 Q24170 KINASE SUPPRESSOR OF R 3.35e+02 665 5 14.3 1390 2 Q69385 EIF4G1 (EUKARYOTIC INI 3.35e+02  
593 5 14.3 966 5 Q24170 KINASE SUPPRESSOR OF R 3.35e+02 666 5 14.3 1404 4 Q43177 CORE PROTEIN. 3.35e+02  
594 5 14.3 972 2 P73273 CATION-TRANSPORTING P- 3.35e+02 667 5 14.3 1409 2 Q52661 F316.24 PROTEIN. 3.35e+02  
595 5 14.3 974 13 Q91735 EPH RECEPTOR TYROSINE 3.35e+02 668 5 14.3 1417 10 Q48697 T19C21.7 PROTEIN. 3.35e+02  
596 5 14.3 985 14 Q67643 INFECTIOUS LARINGOTRAC 3.35e+02 669 5 14.3 1421 10 Q80907 LUSTRIN A. 3.35e+02  
597 5 14.3 989 2 Q59736 2-OXOGLUTARATE DEHYDRO 3.35e+02 670 5 14.3 1428 5 Q44341 HLS-HOMELSS (FRAGMENT 3.35e+02  
598 5 14.3 991 4 Q43335 KIAA0420. 3.35e+02 671 5 14.3 1441 5 Q26453 UDP-GLC:GLYCOPROTEIN G 3.35e+02  
599 5 14.3 991 10 Q81330 F3D13.1 PROTEIN. 3.35e+02 672 5 14.3 1447 3 Q09140 SPIKE PROTEIN. 3.35e+02  
600 5 14.3 991 4 Q43303 KIAA0419. 3.35e+02 673 5 14.3 1453 14 Q91603 GLUCOSYLTRANSFERASE-SI 3.35e+02  
601 5 14.3 995 5 P90897 F58G11.2 PROTEIN. 3.35e+02 674 5 14.3 1455 2 Q69388 GLUCOSYLTRANSFERASE-SI 3.35e+02  
602 5 14.3 998 4 Q43501 SARCO-/ENDOPLASMIC RET 3.35e+02 675 5 14.3 1455 2 Q69397 GLUCOSYLTRANSFERASE-SI 3.35e+02  
603 5 14.3 999 11 Q64518 SARCOENDOPLASMIC RETIC 3.35e+02 676 5 14.3 1455 2 Q69391 LAMININ ALPHA 3B CHAIN 3.35e+02  
604 5 14.3 1001 5 Q20649 CODED FOR BY C. ELEGAN 3.35e+02 677 5 14.3 1486 4 Q14637





```

970 4 11.4 1123 5 043948 PK4 PROTEIN KINASE. 6.37e+03
971 4 11.4 1125 5 027810 DYNEIN HEAVY CHAIN ISO 6.37e+03
972 4 11.4 1165 4 032619 MYELOBLAST KIAA0223 (F 6.37e+03
973 4 11.4 1170 6 028179 THROMBOSPONDIN 1. 6.37e+03
974 4 11.4 1183 2 086064 CYTOTOXIN ASSOCIATED P 6.37e+03
975 4 11.4 1183 3 089965 CHROMOSOME XVI READING 6.37e+03
976 4 11.4 1231 5 046046 1-EVIDENCE-PREDICTED B 6.37e+03
977 4 11.4 1238 2 069947 HYPOTHETICAL 133.5 KD 6.37e+03
978 4 11.4 1268 5 093191 FL3D2.1 PROTEIN. 6.37e+03
979 4 11.4 1273 4 060316 KIAA0570 PROTEIN. 6.37e+03
980 4 11.4 1314 5 027812 DYNEIN HEAVY CHAIN ISO 6.37e+03
981 4 11.4 1337 2 076465 F1337. 6.37e+03
982 4 11.4 1341 3 082821 CHROMOSOME XV READING 6.37e+03
983 4 11.4 1405 5 024341 LARGE FORKED PROTEIN. 6.37e+03
984 4 11.4 1411 2 076701 RHB CORE PROTEIN WITH 6.37e+03
985 4 11.4 1462 2 069389 GLUCOSYLTRANSFERASE-S 6.37e+03
986 4 11.4 1462 2 069392 GLUCOSYLTRANSFERASE-S. 6.37e+03
987 4 11.4 1494 5 020943 SIMILAR TO THE HUMAN M 6.37e+03
988 4 11.4 1544 10 065537 HYPOTHETICAL 173.9 KD 6.37e+03
989 4 11.4 1587 4 015741 PROTEIN DENN. 6.37e+03
990 4 11.4 1770 3 012491 TY2B PROTEIN. 6.37e+03
991 4 11.4 1788 4 060612 SUPERVILLIN. 6.37e+03
992 4 11.4 1849 4 04572 PAC CLONE 248015 FROM 6.37e+03
993 4 11.4 2005 4 034686 KIAA0181 PROTEIN (FRAG 6.37e+03
994 4 11.4 2022 4 034599 MYOSIN-IXB. 6.37e+03
995 4 11.4 2531 5 021980 SIMILAR TO TENASCIN. 6.37e+03
996 4 11.4 2700 5 019689 COSMID F21H11. 6.37e+03
997 4 11.4 2924 5 025733 PFENP1 VARIANT 1 OF ST 6.37e+03
998 4 11.4 3078 5 026031 VARIANT-SPECIFIC SURFA 6.37e+03
999 4 11.4 3394 5 077384 MAL3P6.11 PROTEIN. 6.37e+03
1000 4 11.4 5170 5 017490 CODED FOR BY C. ELEGAN 6.37e+03

```

## ALIGNMENTS

```

RESULT 1
ID O88826 PRELIMINARY; PRT; 913 AA.
AC O88826;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE G0B-5 PROTEIN.
GN G0B-5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE;
RA KOMIYA T., TANIGAWA Y., HIROHASHI S.;
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet
cells in mice."
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AB017156; D1034712;
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 25.7%; Score 9; DB 11; Length 913;
Best Local Similarity 100.0%; Pred.No. 3.45e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 726 CFSRTSSGG 734
|11111111
QY 27 CFSRTSSGG 35

RESULT 2
ID Q00409 PRELIMINARY; PRT; 287 AA.
AC Q00409;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE METHYLVIOLIGEN-REDUCING HYDROGENASE VHU GAMMA SUBUNIT (EC 1.2.-.-)
OE (VHUG).

```

```

GN VHUG.
OS METHANOCOCCUS VOLTAE.
OC ARCHAEA; EUYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PS(DSM1537);
RX MEDLINE; 92293118.
RA HALBOTH S., KLEIN A.;
RT "Methanococcus voltae harbors four gene clusters potentially encoding
two [NifE] and two [Nifese] hydrogenases, each of the cofactor
F420-reducing or F420-non-reducing types."
RL MOL. GEN. GENET. 233:217-224(1992).
RN [2]
RP REVISIONS.
RA PREIFER M., KLEIN A.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- COFACTOR: VHU CONTAINS NI, FE/S, AND FAD COFACTORS.
DR EMBL; X61204; E286614;
KW OXIDOREDUCTASE; NICKEL; IRON-SULFUR; 4FE-4S; FAD; METAL-BINDING;
KW ELECTRON TRANSPORT.
FT METAL 15 15 IRON-SULFUR (BY SIMILARITY).
FT METAL 18 18 IRON-SULFUR (BY SIMILARITY).
FT METAL 234 234 IRON-SULFUR (BY SIMILARITY).
FT METAL 237 237 IRON-SULFUR (BY SIMILARITY).
SQ SEQUENCE 287 AA; 30914 MW; 71642D22 CRC32;

Query Match 20.08; Score 7; DB 1; Length 287;
Best Local Similarity 100.0%; Pred.No. 2.01e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 151 PPRPEIN 157
|11111111
QY 11 PPRPEIN 17

RESULT 3
ID Q21640 PRELIMINARY; PRT; 84 AA.
AC Q21640;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE COSMID R02E12.
GN R02E12.5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSWORTH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN N., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WORLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA LEIMBACH D.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA WATERSTON R.;

```

RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U53337; G1255834; -;  
SQ SEQUENCE 84 AA; 9450 MW; B5D1A6AF CRC32;

Query Match 17.1%; Score 6; DB 5; Length 84;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 FSRTSS 10  
|||||  
Oy 28 FSRTSS 33

RESULT 4  
ID O24602 PRELIMINARY; PRT; 131 AA.

AC O24602;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE ESRIG2 (ESRIG1 PROTEIN).  
GN ESRIG2 OR ESRIG1 OR ESRIG1.  
OS ZEA MAYS (MAIZE).  
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILLOPSIDA; POALES;  
OC POACEAE; ZEA.  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=CV. A188 AND CV. HYBRID HD5XHD7;  
RA OPSAH-FERSTAD H.G., LE DEUNFF E., DUMAS C., ROGOWSKY P.M.;  
RL PLANT J. 12:235-246(1997).  
DR EMBL; X99968; E258942; -;  
DR EMBL; X98497; E258940; -;  
DR EMBL; X98496; E258936; -;  
SQ SEQUENCE 131 AA; 14432 MW; 0F8FF145 CRC32;

Query Match 17.1%; Score 6; DB 10; Length 131;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 112 SRTSSG 117  
|||||  
Oy 29 SRTSSG 34

RESULT 5  
ID P96325 PRELIMINARY; PRT; 170 AA.

AC P96325;  
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE PUTATIVE MINOR FIMBRIAL SUBUNIT.  
GN RALF.  
OS ESCHERICHIA COLI.  
OC PLASMID PRAP.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=83/39.  
RX MEDLINE; 98053968.  
RA ADAMS L.M., SIMMONS C.P., REZMANN L., STRUGNELL R.A.,  
RA ROBINS-BROWNE R.M.;

"Identification and characterization of a k88- and CS31A-like operon of a rabbit enteropathogenic Escherichia coli strain which encodes fimbriae involved in the colonization of rabbit intestine.";  
RL INFECT. IMMUN. 65:5222-5230(1997).  
DR EMBL; U84144; G1850974; -;  
KW PLASMID.

SQ SEQUENCE 170 AA; 18616 MW; 8E5F88C2 CRC32;

Query Match 17.1%; Score 6; DB 2; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 79 RTSSG 84  
|||||  
Oy 30 RTSSG 35

RESULT 6  
ID O57823 PRELIMINARY; PRT; 184 AA.

AC O57823;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE 184AA LONG HYPOTHETICAL PROTEIN.  
GN PH0099.

OS PYROCOCCUS HORIKOSHII.  
OC ARCHAEA; EURYARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCCUS.  
[1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RX MEDLINE; 98344137.

RA KAWARABAYASI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,  
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSIYAMA A., NAGAI Y.,  
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,  
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,  
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,  
RA KIKUCHI H.;

RT "Complete Sequence and Gene Organization of the Genome of a  
Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA RES. 5:55-76(1998).  
DR EMBL; AP000001; D1030111; -;

SQ SEQUENCE 184 AA; 21095 MW; AE8CFFC3 CRC32;

Query Match 17.1%; Score 6; DB 1; Length 184;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 IENDEI 99  
|||||  
Oy 2 IENDEI 7

RESULT 7

ID O60599 PRELIMINARY; PRT; 218 AA.

AC O60599;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
DE PUTATIVE PEROXISOME MICROBODY PROTEIN 175.1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
[1]

RP SEQUENCE FROM N.A.

RA PAGE N.M., BUTLIN D.J., MANYONDA I., BICKNELL A.B., LOWRY P.J.;

RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; AF050199; G2961557; -;

SQ SEQUENCE 218 AA; 24247 MW; 9058FF9A CRC32;

Query Match 17.1%; Score 6; DB 4; Length 218;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 142 FSRTSS 147  
|||||  
Oy 28 FSRTSS 33

RESULT 8

ID P91170 PRELIMINARY; PRT; 249 AA.

AC P91170;

DT 01-MAY-1997 (TREMBLREL. 03, CREATED)

DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE SIMILARITY TO TOXOCARA CANIS SURFACE COAT GLYCOPROTEIN TES-120

DE U39815.  
GN C46H11.9.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL NATURE 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA MILLER N., BRADSHAW H., WAMSLEY P.,  
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA WATERSTON R.,  
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U88314; G1825773; -.  
SQ SEQUENCE 249 AA; 25692 MW; 259347D9 CRC32;

Query Match 17.1%; Score 6; DB 5; Length 249;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 86 SRTSSG 91  
|||||  
QY 29 SRTSSG 34

RESULT 9  
ID O54487 PRELIMINARY; PRT; 263 AA.  
AC O54487;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE SCB PRECURSOR.  
GN SCB.  
OS STAPHYLOCOCCUS CARNOSUS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC STAPHYLOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TM 300;  
RA KRISMER B., MONTENBRUCK I., GOETZ F.,  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U96107; G2735506; -.  
KW SIGNAL.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 263 SCB.  
SQ SEQUENCE 263 AA; 28074 MW; F88E8B33 CRC32;

Query Match 17.1%; Score 6; DB 2; Length 263;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 SRTSSG 158  
|||||

QY 29 SRTSSG 34

RESULT 10  
ID Q91953 PRELIMINARY; PRT; 293 AA.  
AC Q91953;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 32.5 KD PROTEIN.  
OS GALLUS GALLUS (CHICKEN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-LENS FIBERS;  
RX MEDLINE; 96437509.  
RA SAWADA K., AGATA K., EGUCHI G.,  
RT "Characterization of terminally differentiated cell state by  
RT categorizing cDNA clones derived from chicken lens fibers";  
RL INT. J. DEV. BIOL. 40:531-535(1996).  
DR EMBL; D26340; D1005924; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 293 AA; 32499 MW; 3009AAD9 CRC32;

Query Match 17.1%; Score 6; DB 13; Length 293;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 238 SRTSSG 243  
|||||  
QY 30 SRTSSG 35

RESULT 11  
ID Q21090 PRELIMINARY; PRT; 314 AA.  
AC Q21090;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE K01D12.1 PROTEIN.  
GN K01D12.1.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DOBSON R.,  
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL NATURE 368:32-38(1994).  
DR EMBL; 275543; E1347661; -.  
SQ SEQUENCE 314 AA; 36578 MW; 0815863D CRC32;

Query Match 17.1%; Score 6; DB 5; Length 314;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 156 PEINCD 161



```

Oy 14 PEINKD 19
|||||
RESULT 12
ID Q19380 PRELIMINARY; PRT; 376 AA.
AC Q19380;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE COSMID F13B9.
GN F13B9.5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL NATURE 368:32-38(1994).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC STELLYES L.;
RA SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC WATERSTON R.;
RA SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U39853; G1055077;
SQ SEQUENCE 376 AA; 41677 MW; 742EC32B CRC32;

Query Match 17.1%; Score 6; DB 5; Length 376;
Best Local Similarity 100.0%; Pred.No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 152 SRTSSG 157
|||||
Oy 29 SRTSSG 34

RESULT 13
ID Q12076 PRELIMINARY; PRT; 378 AA.
AC Q12076;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SAGA.
GN SAGA.
OS EMERICELLA NIDULANS (ASPERGILLUS NIDULANS).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PLECTOMYCETES; EUROTIALES;
OC TRICHOMACEAE; EMERICELLA.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-A26; TISSUE-MYCELIA;

```

```

RA JONES G.W., HOOLEY P., FARRINGTON S.M., SHAWCROSS S.G., IWANEJKO L.A.,
RA STRIKE P.;
RT "Cloning and characterization of the sagA gene of Aspergillus
RT nidulans.";
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U28804; G3377791;
DR PFAM: PF00036; ehand: 1.
SQ SEQUENCE 378 AA; 42779 MW; 7C3084FB CRC32;

Query Match 17.1%; Score 6; DB 3; Length 378;
Best Local Similarity 100.0%; Pred.No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 189 PEINKD 194
|||||
Oy 14 PEINKD 19

RESULT 14
ID O02236 PRELIMINARY; PRT; 393 AA.
AC O02236;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE E01G4.1 PROTEIN.
GN E01G4.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RN SEQUENCE FROM N.A.
RA SIMS M.;
RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL NATURE 368:32-38(1994).
DR EMBL: Z83223; E1344975;
SQ SEQUENCE 393 AA; 45297 MW; A1847967 CRC32;

Query Match 17.1%; Score 6; DB 5; Length 393;
Best Local Similarity 100.0%; Pred.No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 279 PPRPEI 284
|||||
Oy 11 PPRPEI 16

RESULT 15
ID O27300 PRELIMINARY; PRT; 423 AA.
AC O27300;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HOMOSERINE DEHYDROGENASE.
GN MTH1232.
OS METHANOBACTERIUM THERMOAUTOTROPHICUM.
OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;
OC METHANOBACTERIUM.
RN [1]

```

RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H:  
RX MEDLINE; 98037514.  
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,  
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,  
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,  
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,  
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,  
RA MCDUGALL S., SHIMER G., GOVAL A., PIETROVSKI S., CHURCH G.M.,  
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
deltaH: functional analysis and comparative genomics.";  
RL J. BACTERIOL. 179:7135-7155(1997).  
DR EMBL; AE000890; G2622343; -;  
DR PFAM; PF00742; Homoserine\_dh; 1.  
SQ SEQUENCE 423 AA; 45913 MW; 02BE7C24 CRC32;

Query Match 17.1%; Score 6; DB 1; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 338 FSRTSS 343  
|  
QY 28 FSRTSS 33

Search completed: Sat Aug 28 15:07:15 1999  
Job time : 50 secs.

\*\*\*\*\*  
MPSRCH\_PP protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:16:47 1999; MasPar time 4.50 Seconds  
Tabular output not generated.  
Title: >US-09-049-696-47  
Description: (1-31) from US09049696.ppe  
Perfect Score: 213  
Sequence: 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31  
Scoring table: PAM 150  
Gap 11  
Searched: 170751 seqs, 21266608 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: a-genes35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39  
Statistics: Mean 22.128; Variance 92.385; scale 0.240  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
SUMMARIES  
Result No. Score Query Match Length DB ID Description Pred. No.  
1 139 65.3 228 20 W06548 Human colon specific 6.26e-05  
2 139 65.3 228 29 W46879 Protein sequence enco 6.26e-05  
3 72 33.8 2496 26 W36093 Mutant YLR087c protei 3.93e+01  
4 72 33.8 2942 25 W22050 Saccharomyces cerevis 3.93e+01  
5 71 33.3 252 4 R22997 Yeast proteasome YC7- 4.73e+01  
6 70 32.9 731 32 W44849 S. pneumoniae penicil 5.67e+01  
7 69 32.4 1577 16 R91047 Alpha-D-glucosyltrans 6.80e+01  
8 68 31.9 360 19 R99274 Chemokine receptor K5 8.16e+01  
9 68 31.9 991 16 R80096 Black widow spider 'de 8.16e+01  
10 68 31.9 1214 16 R80097 Black widow spider de 8.16e+01  
11 67 31.5 58 20 R99183 Genetically engineered 9.77e-01  
12 67 31.5 1822 5 R27745 Extracellular factor 9.77e-01  
13 66 31.0 372 28 W44284 N-methylamino acid ox 1.17e+02  
14 66 31.0 737 13 R70186 Arg-gingipain-1 1.17e+02  
15 66 31.0 737 29 W34846 Arg-gingipain high mo 1.17e+02  
16 66 31.0 991 14 R77313 Porphyromonas gingiva 1.17e+02

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

17 66 31.0 1687 17 R96033 P. gingivalis haemagg 1.17e-02  
18 66 31.0 1687 36 W69495 Haemagglutinin protei 1.17e-02  
19 66 31.0 1704 13 R70188 Arg-gingipain-2 prope 1.17e-02  
20 66 31.0 1704 28 W34843 Arg-gingipain high mo 1.17e-02  
21 66 31.0 1706 24 W24786 P. falciparum protei 1.17e-02  
22 66 31.0 2703 13 R70236 P. falciparum Proj3. 1.17e-02  
23 66 31.0 2710 23 W22482 Plasmodium var7. 1.17e-02  
24 66 31.0 3060 23 W22475 Plasmodium var7. 1.17e-02  
25 66 30.5 554 31 W56099 A. Oryzae ATCC20386 c 1.40e+02  
26 65 30.5 843 5 R27744 Extracellular protein 1.40e+02  
27 65 30.5 855 6 R29815 S receptor kinase pro 1.40e+02  
28 64 30.0 854 30 W56309 Clas II S-receptor ki 1.69e-02  
29 63 29.6 217 14 R75460 Mouse antibody H3-3 h 1.99e-02  
30 63 29.6 717 32 W59008 Leucocin A ATP transp 1.99e-02  
31 62 29.1 45 31 W28353 Amino acid sequence o 2.38e-02  
32 62 29.1 307 28 W38334 Rat kidney injury rel 2.38e-02  
33 62 29.1 458 20 W09020 Neuronal nicotinic ac 2.38e-02  
34 62 29.1 474 14 R80507 S. lividans protease 2.38e-02  
35 62 29.1 884 7 R34402 Sequence of low molec 2.38e-02  
36 62 29.1 887 7 R34403 Sequence of low molec 2.38e-02  
37 61 28.6 177 8 R39712 A. Oryzae C6A neutral 2.83e-02  
38 61 28.6 211 22 W20755 H. pylori protein. 2.83e-02  
39 61 28.6 248 1 P91216 Domain VI of mouse la 2.83e-02  
40 61 28.6 257 29 W55462 H. pylori ORF 07ap110 2.83e-02  
41 61 28.6 352 3 R14147 Pre-pro neutral prote 2.83e-02  
42 61 28.6 511 39 W73506 VSV-G protein sequenc 2.83e-02  
43 61 28.6 511 19 W04539 Vesiculovirus glycopr 2.83e-02  
44 61 28.6 1288 29 W55347 H. pylori ORF 14ee19 2.83e-02  
45 61 28.6 1786 35 W50893 Human laminin B1 chai 2.83e-02

ALIGNMENTS

RESULT 1  
ID W06548 standard; Protein; 228 AA.  
AC W06548;  
DT 13-MAR-1997 (first entry)  
DE Human colon specific gene CSG5 polypeptide fragment.  
KW Colon specific gene; CSG5; colon cancer; metastasis; diagnosis;  
KW therapy; antibody; vaccine.  
OS Homo sapiens.  
PN W0639419-A1.  
PD 12-DEC-1996.  
PF 06-JUN-1995; U07289.  
PR 06-JUN-1995; WO-007289.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Yu G;  
DR WPI: 97-043054/04.  
DR N-PSDB: T45884.  
PT Human colon specific genes and their expression products - detection  
PT of which, in non-colon tissue samples, can be used as indication of  
PT colon cancer metastasis  
PS Claim 8: Fig 5; 60pp; English.  
CC Novel polypeptides (W06545-53) are encoded by cDNA clones (see also  
CC T45880-92) corresponding to 13 human colon specific genes  
CC designated CSG1, CSG2, etc., that are primarily expressed in  
CC tissues derived from the colon. Recombinant CSG polypeptides can  
CC be produced in transformed host cells. They are useful diagnostic  
CC markers for colon cancer and for colon cancer metastasis and can  
CC also be used to screen for (ant)agonist cpds. of therapeutic or  
CC diagnostic value. Antibodies raised against the colon-specific  
CC polypeptides may be used to target colon cancer cells or as part  
CC of a colon cancer vaccine.  
SQ Sequence 228 AA;

Query Match 65.3%; Score 139; DB 20; Length 228;

Best Local Similarity 100.0%; Pred. No. 6.26e-05; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vvsryfttydngrysvk 18

QY 14 vvsryfttydngrysvk 31

```

2
RESULT
ID W46879 standard; Protein: 228 AA.
AC W46879;
DT 22-JUN-1998 (first entry)
DE Protein sequence encoded by a colon-specific gene.
KW Colon-specific gene; probe; detection; expression; human;
KW diagnostic assay; colon cancer; antibody; screening.
OS Homo sapiens.
PN US5733748-A.
PD 31-MAR-1998.
PF 06-JUN-1995; 469667.
PR (HUMA-) HUMAN GENOME SCI INC.
PI Rosen C, Yu G;
DR WPI: 98-229823/20.
DR N-PSDB: V16672.
PT Colon-specific nucleic acids - useful as probes for detecting colon
PT cancer micrometastases
PS Claim 1; Fig 5; 51pp; English.
CC W46876-80 and W4682-85 represent proteins encoded by colon-specific
CC genes. The polynucleotides encoding these proteins can be used
CC as probes to detect expression of the corresponding human genes,
CC e.g. in diagnostic assays for detecting micrometastases of colon cancer.
CC Recombinant cells containing the polynucleotides can be used to
CC produce the proteins, in order that antibodies can be raised and
CC used in further screening or diagnostics.
SQ Sequence 228 AA:

Query Match 65.3%; Score 139; DB 29; Length 228;
Best Local Similarity 100.0%; Pred. No. 6.26e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 vvsryfitydngrysvk 18
| | | | | | | | | | | | | | | | | | | |
Qy 14 VYSRYFTTYDTNGRYSVK 31

3
RESULT
ID W36093 standard; Protein: 2496 AA.
AC W36093;
DT 25-MAR-1998 (first entry)
DE Mutant YLR087c protein from cold sensitive yeast strain.
KW Mutant; cold sensitive; yeast; complementation; breadmaking; dough;
KW sugar; leavening agent; brewing; winemaking; truncation.
OS Saccharomyces cerevisiae.
FH Key Location/Qualifiers
FT Misc_difference 527 /note= "encoded by GCR"
FT Misc_difference 618 /note= "encoded by GAY"
FT Misc_difference 640 /note= "encoded by TCY"
FT Misc_difference 1103 /note= "encoded by AGR"
FT Misc_difference 1582 /note= "encoded by GAR"
FT WO9728693-A1.
PN 14-AUG-1997.
PF 07-FEB-1997; F00254.
PR (LESA) LESAFFRE & CIE.
PI Colavizza D, Loiez A, Wadoux I;
DR WPI: 97-414988/38.
DR N-PSDB: T94548.
PT New strains of bread-making yeast with low fermentative activity at
PT low temperature - allows production of doughs that can be stored
PT cold for many hours before final baking
PS Example 4; Page 59-64; 76pp; French.
CC This is the amino acid sequence of a mutant protein (designated YLR087c)
CC isolated from a cold sensitive strain of Saccharomyces cerevisiae. The
CC gene sequence was isolated from the clone Ycp50-10.39 which was able to
CC complement the cold sensitive strain HLI3.2.30. This protein is a

```

```

CC truncated mutation as compared to the wild type sequence. The wild type
CC gene contains a open reading frame of 8874 bases encoding a protein of
CC 2958 amino acids. The mutant gene has an open reading frame of 7488
CC bases encoding a protein of 2496 amino acids. Yeast strains containing
CC the YLR087c gene are used, fresh or dried, in modified breadmaking
CC processes, i.e. in processes where the delay between mixing the dough
CC and baking exceeds 6 hr. Particularly they are used to make French-style
CC bread (no added sugar) or breads with sugar content below 5%. The strains
CC can also be used to make mixed yeast/bacteria acidic leavening agents.
CC The strains can also be used in brewing and winemaking as a
CC cold-sensitive phenotype.
SQ Sequence 2496 AA;

```

```

Query Match 33.8%; Score 72; DB 26; Length 2496;
Best Local Similarity 34.5%; Pred. No. 3.93e+01;
Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 2;

```

```

Db 2092 gagsnpgmlg-ynrlfsayekdfgkftv 2119
| | | | | | | | | | | | | | | | | | | |
Qy 3 GAGADATKDDGVYSRYFTTYDTN-GRYSV 30

4
RESULT
ID W22050 standard; Protein: 2942 AA.
AC W22050;
DT 26-FEB-1998 (first entry)
DE Saccharomyces cerevisiae protein encoded by novel CSF-1 gene.
KW Saccharomyces cerevisiae; cold sensitivity of fermentability; CSF-1;
KW low temperature sensitivity; yeast; bread; ethanol; fermentation.
OS Saccharomyces cerevisiae.
PN WO9724442-A1.
PD 10-JUL-1997.
PF 27-DEC-1996; J03862.
PR 28-DEC-1995; JP-343700.
PA (KYOM) KYOMA HAKKO KOGYO KK.
PI Kawasaki H, Kikuchi Y, Ouchi K, Tokai M;
DR WPI: 97-363678/33.
DR N-PSDB: T74989.
PT Protein for complementing low temperature sensitivity in
PT fermentation - can be inactivated in yeast and used for producing
PT bread and ethanol
PS Claim 3; Pages 21-40; 48pp; Japanese.
CC This is a protein capable of complementing a mutation showing low
CC temperature sensitivity in fermentation. This is encoded by a novel
CC genomic DNA CSF-1 (cold sensitivity of fermentability). The Saccharomyces
CC cerevisiae YHK 1243 has an inactivated gene encoding this new protein or
CC has the novel DNA sequence contained in its chromosome. This protein can
CC be used in dough to produce bread, and to produce ethanol.
SQ Sequence 2942 AA;

```

```

Query Match 33.8%; Score 72; DB 25; Length 2942;
Best Local Similarity 34.5%; Pred. No. 3.93e+01;
Matches 10; Conservative 11; Mismatches 6; Indels 2; Gaps 2;

```

```

Db 2092 gagsnpgmlg-ynrlfsayekdfgkftv 2119
| | | | | | | | | | | | | | | | | | | |
Qy 3 GAGADATKDDGVYSRYFTTYDTN-GRYSV 30

5
RESULT
ID R22997 standard; Protein: 252 AA.
AC R22997;
DT 08-OCT-1992 (first entry)
DE Yeast proteasome YC7-alpha subunit.
KW Proteasome; proteolytic.
OS Saccharomyces cerevisiae.
PN J04077497-A.
PD 11-MAR-1992.
PF 19-JUL-1990; 192113.
PR 19-JUL-1990; JP-192113.
PA (SAKA) OTSUKA PHARM KK.
DR WPI: 92-136766/17.
DR N-PSDB: Q23893.

```

PT Yeast proteasome with specific aminoacid sequence - has unique  
 PT proteolytic activity, used for decomposition and synthesis of  
 PT protein  
 PS Claim 3; page 2; 26pp; Japanese.  
 CC The yeast proteasome has unique proteolytic activity and  
 CC can be used in the synthesis and decomposition of protein.  
 SQ Sequence 252 AA;

Query Match 33.3%; Score 71; DB 4; Length 252;

Best Local Similarity 39.3%; Pred. No. 4.73e+01;

Matches 11; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db 2 sgaaasaaq-ydrhifsfpegrlyqv 28

QY 4 AGADATKDDGVYSRYFTTYDTNGR-YSV 30

# RESULT 6

ID W44849 standard; Protein; 731 AA.

AC W44849;

DT 21-AUG-1998 (first entry)

DE S. pneumoniae penicillin-binding protein PBP-Nv.

KW Penicillin-binding protein; PBP-Nv; transglycolase; antibacterial agent;

OS trans-peptidation activity; infection; therapy.

PN Streptococcus pneumoniae.

PD EP-837132-A2.

PF 17-OCT-1997; 308288.

PR 17-OCT-1996; US-731716.

PA (ELIL ) LILLY & CO ELI.

PI Hoskins JA, Jaskunas SR, Norris FH, Rostock PR,

PI Zhao G;

DR WPI; 98-219114/20.

DR N-PSDB; V19373.

PT Recombinant Streptococcus pneumoniae penicillin-binding protein -

PT useful in screening assays for antibacterial agents

PS Claim 1; Page 19-21; 28pp; English.

CC This sequence is the Streptococcus pneumoniae penicillin-binding

CC protein, designated PBP-Nv, of the invention. The protein is useful in

CC screening assays for compounds that bind to PBP-Nv or inhibit the

CC transglycolase or trans-peptidation activity of PBP-Nv. Such compounds

CC would be useful as antibacterial agents for treating S. pneumoniae

CC infections.

SQ Sequence 731 AA;

Query Match 32.9%; Score 70; DB 32; Length 731;

Best Local Similarity 47.4%; Pred. No. 5.67e+01;

Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Db 137 kndginygrfflaivtagr 155

QY 10 KDDGV-YSRYFTTYDTNGR 27

# RESULT 7

ID R91047 standard; Protein; 1577 AA.

AC R91047;

DT 22-MAY-1996 (first entry)

DE Alpha-D-glucosyltransferase.

KW sucrose; transgenic plant; cloning; Escherichia coli;

KW phage lambda-C13; vector; plasmid pGS501; plasmid pGS502;

KW gene transfer; crop improvement; storage carbohydrate; pasture;

KW feedstuff; senescence; dextran; binder; food; pharmaceutical.

OS Streptococcus salivarius strain ATCC 25975.

PN WQ9606173-A1.

PD 29-FEB-1996.

PF 24-AUG-1995; AU0527.

PR 24-AUG-1994; AU-007643.

PA (GIFF/) GIFFARD P M.

PA (JACO/) JACQUES N A.

PA (SIMP/) SIMPSON C L.

PI Giffard PM, Jacques NA, Simpson CL;

DR WPI; 96-151376/15.

DR N-PSDB; T13139.

PT Plants con'g. new bacterial DNA encoding glucosyl transferase

PT activity - retain higher levels of stored carbohydrate(s) in a form

PS readily digestible by ruminants

PS Claim 4; Page 16-20; 31pp; English.

CC The sequence represents an alpha-D-glucosyltransferase from

CC Streptococcus salivarius. The enzyme is primer-independent, and

CC produces soluble glucan from sucrose. A gene encoding the enzyme

CC may be cloned and expressed in Escherichia coli using a subclone

CC of phage lambda-C13, e.g. plasmid pGS501 or plasmid pGS502. The

CC DNA may also be expressed in a transgenic plant, to improve the

CC level of stored carbohydrate in a pasture plant which normally

CC contains low levels, or to prevent degradation of stored carbohydrate

CC during plant senescence. Dextran may be isolated from the plant, for

CC use as a food binder or pharmaceutical additive. Primer independence

CC ensures that the enzyme will be functional in plants. The glucan is

CC poorly degraded in plants but easily degraded by bacteria in the rumen

CC of grazing livestock.

SQ Sequence 1577 AA;

Query Match 32.4%; Score 69; DB 16; Length 1577;

Best Local Similarity 36.4%; Pred. No. 6.80e+01;

Matches 8; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Db 317 adaipnnsiyavynqaydtssk 338

QY 6 ADATKDDGVYSRYFTTYDTNGR 27

# RESULT 8

ID R99274 standard; Protein; 360 AA.

AC R99274;

DT 09-DEC-1996 (first entry)

DE Chemokine receptor K5.5.

KW Chemokine receptor K5.5; MIP-1-alpha; RANTES; MCP-1; allergy;

KW atheroma; HIV; AIDS; graft rejection; stem cell.

OS Homo sapiens.

PN WQ9623088-A1.

PD 01-AUG-1996.

PR 24-JAN-1996; G00143.

PR 27-JAN-1995; GB-001683.

PA (GLAX ) GLAXO GROUP LTD.

PI Power CA, Wells TNC;

DR WPI; 96-362692/36.

DR N-PSDB; T35277.

PT Chemokine receptor which binds MIP-1-alpha, RANTES and/or MCP-1 -

PT useful in screening for agents to treat asthma, hay fever, eczema,

PT allergies, atopic dermatitis, rhinitis or conjunctivitis.

PS Claim 1; Fig 3; 47pp; English.

CC Human chemokine receptor K5.5 (R99274) binds MIP-1-alpha, RANTES

CC and/or MCP-1. Its amino acid sequence was deduced from a cDNA

CC clone (T35277) obt'd. from a human spleen lambda gt11 cDNA

CC library. Recombinant chemokine receptor K5.5 can be expressed

CC in transformed host cells. It can be used to screen for agents

CC which act as antagonists to MCP-1, MIP-1-alpha and/or RANTES.

CC Such agents may be useful in treating allergies, atheromas and

CC diseases mediated by viruses, such as AIDS. They can also be

CC used to prevent graft rejection or to protect stem cells from

CC the effects of chemotherapy.

SQ Sequence 360 AA;

Query Match 31.9%; Score 68; DB 19; Length 360;

Best Local Similarity 42.1%; Pred. No. 8.16e+01;

Matches 8; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Db 7 adtldesysnyyl-yes 24

QY 6 ADATKDDGVYSRYFTTYDT 24

# RESULT 9

ID R80096 standard; Protein; 991 AA.

```

Query Match      31.98; Score 68; DB 16; Length 1214;
Best Local Similarity 50.08; Pred. No. 8.fcce-01;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 391 dgvevryavdqeskgmys 408
QY 12 DGVSRSYFTTYDNGRYS 29
    ||| ||| :|||

```

RESULT	12	
ID	R27745	standard; Protein; 1822 AA.
AC	R27745;	
DT	04-MAR-1993	(first entry)
DE	Extracellular	factor related protein.
KW	EF*;	detection; prevention; screening; diagnostic.
OS	Streptococcus suis	type II (non-pathogenic).
EH	Key	Location/Qualifiers
FT	peptide	1..46
FT	peptide	/note= "signal peptide"
FT	peptide	47..1822
FT	region	/note= "mature peptide"
FT	region	858..861
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	934..937
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	999..1002
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	1075..1078
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	1264..1267
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	1362..1365
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	1438..1441
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"
FT	region	/note= "repetitive Asn-Pro-Asn-Leu sequence"



CC KDVTY (V); RMFMNYPGR YTPVEEKONG (VI); TFAGFEDTYK RMFMNYPGR (VII);  
CC DYTIVYRDG TRIKEGLTAT TFEEDGVATG NMEYCVCKY TAGVSPKVC (VIII); YTYTVYRDGT  
CC KIKEGLTATTF EEDG (IX); RDGTIKIEGL TATTFEEDGV ATGN (X); KIKEGLTATT  
CC FEEDGVATGN HEY (XI); KWDAPNGTGN PNP PNPNN PGTTLSE (XII); and  
CC YTPVEEKONG RMIVIVAKKY (XIII). They are used in vaccines to protect  
CC animals, including humans, from gingivitis and/or periodontal  
CC diseases.  
SQ Sequence 737 AA;

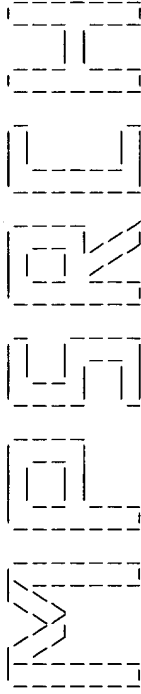
Query Match 31.0%; Score 66; DB 29; Length 737;  
Best Local Similarity 38.9%; Pred. No. 1.17e+02;  
Matches 7; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Db 213 edtykrmfmyep-gryt 229  
Qy 12 DGVYSRYFTTYDTNGRYS 29

Search completed: Mon Aug 23 13:17:16 1999  
Job time : 29 secs.



\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:16:13 1999; MasPar time 5.17 Seconds  
Tabular output not generated. 240.308 Million cell updates/sec

Title: >US-09-049-696-47  
Description: (1-31) from US09049696.pep  
Perfect Score: 213  
Sequence: 1 DNGAGADATKDDGYVSRYFTYDINGRYSVK 31

Scoring table: PAM 150  
Gap 11  
Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 31.785; Variance 59.437; scale 0.535

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	194	91.1	913	3	gob-5 protein - Mouse	1.09e-21
2	164	77.0	794	2	hypothetical protein	7.12e-16
3	164	77.0	820	2	Lu-ECAM-1 protein - b	7.12e-16
4	164	77.0	905	2	Lu-ECAM-1 protein - b	7.12e-16
5	80	37.6	348	1	gene I protein - phag	3.04e-01
6	80	37.6	348	1	gene I protein - phag	3.04e-01
7	80	37.6	348	1	gene I protein - phag	3.04e-01
8	75	35.2	137	2	hypothetical protein	1.63e+00
9	75	35.2	245	2	hypothetical protein	1.63e+00
10	74	34.7	593	2	surface-layer glycopr	2.26e+00
11	74	34.7	593	2	surface-layer glycopr	2.26e+00
12	74	34.7	885	2	acetylglutamate kinas	2.26e+00
13	72	33.8	121	2	chorion protein s15 -	4.32e+00
14	72	33.8	2958	2	probable membrane pro	4.32e+00
15	71	33.3	57	2	protein kinase AK5 (E	5.94e+00
16	71	33.3	252	1	multicatalytic endope	5.94e+00
17	70	32.9	120	2	tungsten formylmethan	8.16e+00
18	70	32.9	245	2	cytosine methyltransf	8.16e+00
19	70	32.9	823	2	hypothetical protein	8.16e+00
20	69	32.4	336	2	HsdS1A protein - Myco	1.12e+01
21	69	32.4	372	2	probable sarcosine ox	1.12e+01
22	69	32.4	401	2	HsdS1A protein - Myco	1.12e+01
23	68	31.9	169	2	hypothetical protein	1.53e+01

24	68	31.9	360	2	A57160	chemokine (C-C) recep	1.53e+01
25	68	31.9	394	2	D71379	probable sensory tran	1.53e+01
26	67	31.5	351	2	B24356	replication-associate	2.07e+01
27	67	31.5	351	2	S49387	replication-associate	2.07e+01
28	67	31.5	403	2	C69213	protein-export membra	2.07e+01
29	66	31.0	57	2	S66314	protein kinase AK1 (E	2.81e+01
30	66	31.0	125	2	H64637	probable translation	2.81e+01
31	66	31.0	272	2	D70312	hypothetical protein	2.81e+01
32	66	31.0	290	2	D69518	conserved hypothetical	2.81e+01
33	66	31.0	331	2	S45584	transcription activat	2.81e+01
34	66	31.0	360	2	JC4587	chemokine (C-C) recep	2.81e+01
35	66	31.0	465	2	F65207	hydH protein - Escher	2.81e+01
36	66	31.0	530	2	S33673	probable protein kina	2.81e+01
37	66	31.0	543	2	C64334	hypothetical protein	2.81e+01
38	66	31.0	843	2	S33442	EF protein - Streptoc	2.81e+01
39	66	31.0	991	2	I40229	arginyl endopeptidase	2.81e+01
40	66	31.0	1104	2	A36866	microbial collagenase	2.81e+01
41	66	31.0	1146	2	A55532	myosin-heavy-chain ki	2.81e+01
42	66	31.0	1526	2	S49763	gingipain R (EC 3.4.2	2.81e+01
43	66	31.0	1662	2	T01893	hypothetical protein	2.81e+01
44	66	31.0	1704	2	A55426	gingipain R (EC 3.4.2	2.81e+01
45	66	31.0	1822	2	S33441	EF protein - Streptoc	2.81e+01

ALIGNMENTS

RESULT 1 JG0168 #type complete  
ENTRY gob-5 protein - Mouse  
TITLE #formal\_name Mus musculus #common\_name house mouse  
ORGANISM 18-Mar-1999 #sequence\_revision 18-Mar-1999 #text\_change  
DATE 18-Mar-1999  
ACCESSIONS JG0168  
REFERENCE JG0168  
#authors Komiyama, T.; Tanigawa, Y.; Hirohashi, S.  
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351  
#title Cloning and identification of the gene gob-5, which is  
#accession JG0168 expressed in intestinal goblet cells in mice.  
#status preliminary  
#residues 1-913 #label KOM  
#cross-references DBJ:AB016592  
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 91.1%; Score 194; DB 3; Length 913;  
Best Local Similarity 83.9%; Pred. No. 1.09e-21;  
Matches 26; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 645 DNGAGADATKDDGYVSRYFTYDANGRYSVK 675  
QY 1 DNGAGADATKDDGYVSRYFTYDINGRYSVK 31  
|||||||:|||||||:|||||||:|||||||:|||||||  
RESULT 2 T02171 #type complete  
ENTRY hypothetical protein Lu-ECAM-1 - bovine  
TITLE #formal\_name Bos primigenius taurus #common\_name cattle  
ORGANISM 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change  
DATE 05-Mar-1999  
ACCESSIONS T02171  
REFERENCE Z14590  
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;  
Levine, R.; Goodwin, A.; Pauli, B.O.

#submission submitted to the EMBL Data Library, April 1997  
#description Cloning and characterization of Lu-ECAM-1 suggest it is an  
endothelial chloride channel.

#accession T02171  
#status preliminary; translated from GB/EMBL/DBJ  
#molecule\_type mRNA  
#residues 1-794 #label ELB  
#cross-references EMBL:AF001262; NID:g2623764; PID:g2623765  
#experimental\_source lung  
SUMMARY #length 794 #molecular-weight 88509 #checksum 6853

```
Query Match          77.0%; Score 164; DB 2; Length 794;
Best Local Similarity 71.0%; Pred. No. 7.12e-16;
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGRDTVKNDCIYRYFTDYXNGRYSLK 682
||||| |: |:|:||||| |:| |:| |:| |:|
QY 1 DNGAGADATKDDGVYSRYFTTYDINGRYSVK 31

RESULT 3
ENTRY T02152 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02152
REFERENCE Z14590
#authors Eble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Paull, B.U.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel.
#accession T02152
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-820 #label ELB
#cross-references EMBL:AF001263; NID:g2623766; PID:g2623767
#experimental_source lung
#length 820 #molecular-weight 91464 #checksum 3104
SUMMARY

Query Match          77.0%; Score 164; DB 2; Length 820;
Best Local Similarity 71.0%; Pred. No. 7.12e-16;
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGRDTVKNDCIYRYFTDYXNGRYSLK 682
||||| |: |:|:||||| |:| |:| |:| |:|
QY 1 DNGAGADATKDDGVYSRYFTTYDINGRYSVK 31

RESULT 4
ENTRY T02205 #type complete
TITLE Lu-ECAM-1 protein - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02205
REFERENCE Z14590
#authors Eble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Paull, B.U.
#description submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel.
#accession T02205
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-905 #label ELB
#cross-references EMBL:AF001261; NID:g2623762; PID:g2623763
#experimental_source lung
#length 905 #molecular-weight 101005 #checksum 5557
SUMMARY

Query Match          77.0%; Score 164; DB 2; Length 905;
Best Local Similarity 71.0%; Pred. No. 7.12e-16;
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGRDTVKNDCIYRYFTDYXNGRYSLK 682
||||| |: |:|:||||| |:| |:| |:| |:|
QY 1 DNGAGADATKDDGVYSRYFTTYDINGRYSVK 31

RESULT 5
ENTRY Z1BPFI #type complete
TITLE gene I protein - phase fl
```

```
ORGANISM #formal_name phase fl
DATE 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change
ACCESSIONS C04262; A04262
REFERENCE A91490
#authors Beck, E.; Zink, B.
#journal Gene (1981) 16:35-58
#title Nucleotide sequence and genome organisation of filamentous
bacteriophages fl and fd.
#cross-references MUID:82211801
#accession C04262
#molecule_type DNA
#residues 1-348 #label BEC
#cross-references GB:V00606; GB:J02449; GB:M10881; NID:g14974;
PID:g14982
COMMENT The exact function of this protein is not known although it may be
involved in phage assembly.
GENETICS
#gene I
CLASSIFICATION #superfamily filamentous phage gene I protein
KEYWORDS phage maturation
SUMMARY #length 348 #molecular-weight 39527 #checksum 2408

Query Match          37.6%; Score 80; DB 1; Length 348;
Best Local Similarity 61.1%; Pred. No. 3.04e-01;
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 215 DSGVYS-VLTPYLSHGRY 231
||||| |:| |:| |:| |:|
QY 11 DSGVYSRYFTTYDTNGRY 28

RESULT 6
ENTRY Z1BPFD #type complete
TITLE gene I protein - phase fd (strain 478, Heidelberg)
ORGANISM #formal_name phase fd
DATE 30-Sep-1980 #sequence_revision 30-Sep-1980 #text_change
ACCESSIONS A04262
REFERENCE A93690
#authors Beck, E.; Sommer, R.; Auerwald, E.A.; Kurz, C.; Zink, B.;
Osterburg, G.; Schaller, H.; Sugimoto, K.; Sugisaki, H.;
Okamoto, T.; Takanami, M.
#journal Nucleic Acids Res. (1978) 5:4495-4503
#title Nucleotide sequence of bacteriophage fd DNA.
#cross-references MUID:79136480
#accession A04262
#molecule_type DNA
#residues 1-348 #label BEC
#cross-references GB:V00602; GB:J02451; GB:M10731; GB:M10767;
GB:M21666; GB:M21667; GB:M21668; GB:M21669;
GB:M21670; GB:M25198; NID:g14931; PID:g14940
COMMENT The exact function of this protein is not known although it may be
involved in phage assembly.
GENETICS
#gene I
CLASSIFICATION #superfamily filamentous phage gene I protein
KEYWORDS phage maturation
SUMMARY #length 348 #molecular-weight 39536 #checksum 2767

Query Match          37.6%; Score 80; DB 1; Length 348;
Best Local Similarity 61.1%; Pred. No. 3.04e-01;
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 215 DSGVYS-VLTPYLSHGRY 231
||||| |:| |:| |:| |:|
QY 11 DSGVYSRYFTTYDTNGRY 28

RESULT 7
ENTRY Z1BPM3 #type complete
TITLE gene I protein - phase M13
ORGANISM #formal_name phase M13
```

```

DATE          31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change
                23-Jan-1998
ACCESSIONS    B04262; A04262
REFERENCE     A91470
#authors      van Wezenbeek, P.M.G.F.; Hulsebos, T.J.M.; Schoenmakers,
                J.G.G.
#journal      Gene (1980) 11:129-148
#title        Nucleotide sequence of the filamentous bacteriophage M13 DNA
                genome: comparison with phage fd.
#cross-references MUID:81067903
#molecule_type DNA
#accession    B04262
#residues     1-348 #label VAN
#cross-references GB:V00604; GB:J02461; GB:M10377; NID:g14959;
                PID:g14968
COMMENT       The exact function of this protein is not known although it may be
                involved in phage assembly.
GENETICS
#gene         I
CLASSIFICATION #superfamily filamentous phage gene I protein
KEYWORDS      phage maturation
SUMMARY       #length 348 #molecular-weight 39550 #checksum 2240
                37.6%; Score 80; DB 1; Length 348;
Query Match   Best Local Similarity 61.1%; Pred. No. 3.04e-01;
Matches       11: Conservative 3; Mismatches 3; Indels 1; Gaps 1;
Db            215 DSGVYS-YLPYLSHGRY 231
                I : I I I I I : I I I I
QY            11 DGGVYSRYFTTYDNGRY 28

RESULT        8
ENTRY         G70215 #type complete
TITLE         hypothetical protein BBA71 - Lyme disease spirochete plasmid
                A/lp54
ORGANISM      #formal_name Borrelia burgdorferi #common_name Lyme disease
                spirochete
DATE          13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
ACCESSIONS    G70215
REFERENCE     A70100
#authors      Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
                Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
                Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
                J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
                Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
                Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
                J.; Utterback, T.; Watthey, L.; McDonald, L.; Artlich, P.;
                Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
                K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
#journal      Nature (1997) 390:580-586
#title        Genomic sequence of a Lyme disease spirochaete, Borrelia
                burgdorferi.
#cross-references MUID:98065943
#accession    G70215
#status       preliminary; nucleic acid sequence not shown;
                translation not shown
#molecule_type DNA
#residues     1-137 #label KLE
#cross-references GB:AF00790; NID:g2690224; PID:g2690288; TIGR:BBA71
#experimental_source strain B31
GENETICS
#genome       plasmid
SUMMARY       #length 137 #molecular-weight 16412 #checksum 9880
                35.2%; Score 75; DB 2; Length 137;
Query Match   Best Local Similarity 28.6%; Pred. No. 1.63e-00;
Matches       6: Conservative 6; Mismatches 9; Indels 0; Gaps 0;
Db            104 KNEDYKKYIYAINENQI 124
                I : I : I : I : I :
QY            10 KDDGVYSRYFTTYDNGRYSV 30

```

```

RESULT        9
ENTRY         E59511 #type complete
TITLE         hypothetical protein AF2093 - Archaeoglobus fulgidus
ORGANISM      #formal_name Archaeoglobus fulgidus
DATE          05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
                05-Jun-1998
ACCESSIONS    E59511
REFERENCE     A59250
#authors      Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
                K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.;
                Peterson, J.D.; Richardson, D.L.; Kerlavage, A.R.; Graham,
                D.E.; Kyrpides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
                Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
                Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.;
                Peterson, S.; Reich, C.I.; McNeill, L.K.; Badger, J.H.;
                Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
                J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
                T.; Artlich, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.;
                D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.;
                Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
                C.R.; Venter, J.C.
#journal      Nature (1997) 390:364-370
#title        The complete genome sequence of the hyperthermophilic,
                sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references MUID:98049343
#accession    E59511
#status       preliminary; nucleic acid sequence not shown;
                translation not shown
#molecule_type DNA
#residues     1-245 #label KLE
#cross-references GB:AE000959; GB:AE000782; NID:g2689282; PID:g2648444;
                TIGR:AF2093
SUMMARY       #length 245 #molecular-weight 28593 #checksum 9803
                35.2%; Score 75; DB 2; Length 245;
Query Match   Best Local Similarity 46.7%; Pred. No. 1.63e-00;
Matches       7: Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Db            114 VTDRFFEGYEAGRY 128
                I : I : I : I : I :
QY            14 VYSRYFTTYDNGRY 28

RESULT        10
ENTRY         S16225 #type complete
TITLE         surface-layer glycoprotein precursor - Methanothermus
                fervidus
ORGANISM      #formal_name Methanothermus fervidus
DATE          22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change
ACCESSIONS    S16225; S21873
REFERENCE     S16225
#authors      Broeckl, G.; Behr, M.; Fabry, S.; Hensel, R.; Kaudewitz, H.;
                Biendl, E.; Koenig, H.
#journal      Eur. J. Biochem. (1991) 199:147-152
#title        Analysis and nucleotide sequence of the genes encoding the
                surface-layer glycoproteins of the hyperthermophilic
                methanogen Methanothermus fervidus and Methanothermus
                sociabilis.
#cross-references MUID:91293115
#accession    S16225
#molecule_type DNA
#residues     1-593 #label BRO
#cross-references EMBL:X58297; NID:g44281; PID:g809714
GENETICS
#gene         sigA
#start_codon GTG
KEYWORDS      glycoprotein
FEATURE       1-22
                23-593
                #domain signal sequence #status predicted #label sig\
                #product surface-layer glycoprotein #status predicted

```

```
#label MAT
SUMMARY      #length 593 #molecular-weight 65481 #checksum 7421
Query Match  34.7%; Score 74; DB 2; Length 593;
Best Local Similarity 35.3%; Pred. No. 2.26e+00;
Matches      6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 295 ASRNKKIYGNFYNYET 311
QY 8 ATKDDGVYSRYFTTYDT 24
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

RESULT 11
ENTRY   S16375 #type complete
TITLE   surface-layer glycoprotein precursor - Methanothermus
        sociabilis
ORGANISM #formal_name Methanothermus sociabilis
DATE     22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change
        13-Sep-1998
ACCESSIONS S16375; S26144; S26098; S21874
REFERENCE  S16225
#authors  Broeckl, G.; Behr, M.; Fabry, S.; Hensel, R.; Kaudewitz, H.;
        Biendl, G.; Koenig, H.
#journal  Eur. J. Biochem. (1991) 199:147-152
#title    Analysis and nucleotide sequence of the genes encoding the
        surface-layer glycoproteins of the hyperthermophilic
        methanogens Methanothermus fervidus and Methanothermus
        sociabilis.
#cross-references MUID:91293115
#accession S16375
#molecule_type DNA
#residues  1-593 #label BRO
#cross-references EMBL:X58296
#accession S26144
#molecule_type protein
#residues  23-42 #label BRO2
REFERENCE  S21873
#authors  Broeckl, G.
#submission submitted to the EMBL Data Library, March 1991
#accession S26098
#molecule_type DNA
#residues  1-256,'I',258,'V',260-593 #label BRO1
#cross-references EMBL:X58296; NID:944546; PID:9809717
GENETICS
#gene      slgA
#start_codon GTG
#keywords  glycoprotein
FEATURE
1-22      #domain signal sequence #status predicted #label SIG\
23-593    #product surface-layer glycoprotein #status experimental
#label MAT
SUMMARY    #length 593 #molecular-weight 65503 #checksum 8058
Query Match  34.7%; Score 74; DB 2; Length 593;
Best Local Similarity 35.3%; Pred. No. 2.26e+00;
Matches      6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 295 ASRNKKIYGNFYNYET 311
QY 8 ATKDDGVYSRYFTTYDT 24
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

RESULT 12
ENTRY   S22389 #type complete
TITLE   acetylglutamate kinase (EC 2.7.2.8) /
        N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) -
        fission yeast (Schizosaccharomyces pombe)
ORGANISM #formal_name Schizosaccharomyces pombe
DATE     30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
        09-Sep-1997
ACCESSIONS S22389; S22174
REFERENCE  S22389
#authors  van Huffel, C.; Dubois, E.; Messenguy, F.
```

```
#journal  Eur. J. Biochem. (1992) 205:33-43
#title    Cloning and sequencing of arg3 and arg11 genes of
        Schizosaccharomyces pombe on a 10-kb DNA fragment.
        Heterologous expression and mitochondrial targeting of
        their translation products.
#cross-references MUID:92209520
#accession S22389
#molecule_type DNA
#residues  1-885 #label VAN1
#cross-references EMBL:X63576
REFERENCE  S22174
#authors  van Huffel, C.H.R.
#submission submitted to the EMBL Data Library, December 1991
#accession S22174
#molecule_type DNA
#residues  1-260,'D',262-344,'P',346-885 #label VAN2
#cross-references EMBL:X63576; NID:94905; PID:94906
GENETICS
#gene      arg11
#map_position 1L
#keywords  mitochondrion; oxidoreductase; phosphotransferase
SUMMARY    #length 885 #molecular-weight 97704 #checksum 5332
Query Match  34.7%; Score 74; DB 2; Length 885;
Best Local Similarity 22.2%; Pred. No. 2.26e+00;
Matches      6; Conservative 12; Mismatches 8; Indels 1; Gaps 1;

Db 70 SSIGSRREVQVLYRFTSFQAF 95
QY 4 AGADATKDDGVYSRYFTTYDNGRYSV 30
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

RESULT 13
ENTRY   S06615 #type complete
TITLE    chorion protein sl5 - fruit fly (Drosophila subobscura)
ORGANISM #formal_name Drosophila subobscura
DATE     07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
        24-Sep-1998
ACCESSIONS S06615
REFERENCE  S06612
#authors  Martinez-Cruzado, J.C.; Swimmer, C.; Fenerjian, M.G.;
        Kafatos, F.C.
#journal  Genetics (1988) 119:663-677
#title    Evolution of the autosomal chorion locus in Drosophila. I.
        General organization of the locus and sequence comparisons
        of genes sl5 and sl9 in evolutionarily distant species.
#cross-references MUID:88297142
#accession S06615
#molecule_type DNA
#residues  1-121 #label MAR
#cross-references GB:X53423; EMBL:X12637; NID:99113; PID:99115
GENETICS
#gene      sl5-1
#cross-references FlyBase:FBgn0012933
#map_position J 28A
#introns  4/3
SUMMARY    #length 121 #molecular-weight 12530 #checksum 5143
Query Match  33.8%; Score 72; DB 2; Length 121;
Best Local Similarity 35.7%; Pred. No. 4.32e+00;
Matches      10; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

Db 77 AASAAAIPGYSQYAIPTYEIDGSYN 104
QY 3 GAGADATKDDGVYSRY-FTTYDTNGRYS 29
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

RESULT 14
ENTRY   S64921 #type complete
TITLE    probable membrane protein YLR087c - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES hypothetical protein L2506
ORGANISM #formal_name Saccharomyces cerevisiae
```

```

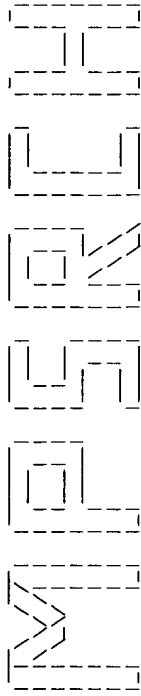
Db      31 ANTKRIYGTG-YMSPYAHGKYSVK 56
      | | | | |
      | | | | |
QY      6 ADATKDDGVYSRYFTT-YDINGRYSVK 31

Search completed: Mon Aug 23 13:16:28 1999
Job time : 15 secs.

```

**This Page Blank (uspto)**

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch pp protein - protein database search, using Smith-Waterman algorithm

```
Run on:      Mon Aug 23 13:15:02 1999;  MasPar time 3.72 Seconds
           235.539 Million cell updates/sec
Tabular output not generated.
```

```
>US-09-049-696-47
Description:
Perfect Score: 213
Sequence: 1 DNGAGADATKDDGVYSRYFTYDTNGRYSVK 31
```

Scoring table: PAM 150  
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0% Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 32.535; Variance 54.091; scale 0.601

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description	Pred. No.
1	173	81.2	903	1	ECLC_BOVIN	EPITHELIAL CHLORIDE CH	5.16e+20
2	80	37.6	348	1	VGL_BPF1	GENE I PROTEIN.	8.05e-02
3	80	37.6	348	1	VGL_BPM13	GENE I PROTEIN.	8.05e-02
4	80	37.6	348	1	VGL_BPF2	GENE I PROTEIN.	8.05e-02
5	74	34.7	593	1	CSG_METTE	CELL SURFACE GLYCOPROT	7.32e-01
6	74	34.7	593	1	CSG_METSC	CELL SURFACE GLYCOPROT	7.32e-01
7	74	34.7	885	1	ARS6_SCHPO	CELL SURFACE GLYCOPROT	7.32e-01
8	73	34.3	369	1	HIS8_ZYMO	ARISTIDINOL-PRECURSOR	1.05e+00
9	72	33.8	121	1	HIS15_DROSU	HISTIDINOL-PREPHATE A	1.05e+00
10	71	33.3	252	1	PRCI_YEAST	CHORION PROTEIN S15.	1.49e+00
11	70	32.9	823	1	YHX2_YEAST	PTHEASOME COMPONENT C	2.12e+00
12	69	32.4	372	1	SAOX_ECOLI	HYPOTHETICAL 96.8 KD P	3.00e+00
13	68	31.9	360	1	KR4_HUMAN	PUTATIVE SARCOSINE OXI	4.24e+00
14	67	31.5	436	1	VE2_HPV22	C-C CHEMOKINE RECEPTOR	5.96e+00
15	66	31.0	125	1	Y944_HELPY	REGULATORY PROTEIN E2.	8.35e+00
16	66	31.0	360	1	KR4_MOUSE	HYPOTHETICAL PROTEIN H	1.17e+01
17	66	31.0	463	1	LYDC_ECOLI	C-C CHEMOKINE RECEPTOR	1.17e+01
18	66	31.0	475	1	HACH_LACAC	SENSOR PROTEIN HYDH (E	1.17e+01
19	66	31.0	530	1	PILS_PSEA	6-PHOSPHO-BETA-GALACTO	1.17e+01
20	66	31.0	543	1	Y274_METJA	SENSOR PROTEIN PILS (E	1.17e+01
21	66	31.0	991	1	CPGI_PORGI	HYPOTHETICAL PROTEIN M	1.17e+01
22	66	31.0	1104	1	COLA_CLOPE	GINGIPAIN R1 PRECURSOR	1.17e+01
23	66	31.0	1146	1	KMHA_DICDI	MYOSIN HEAVY CHAIN	1.17e+01

```
FT CARBOHYD 515 515 POTENTIAL.
FT CARBOHYD 688 688 POTENTIAL.
FT CARBOHYD 811 811 POTENTIAL.
FT CARBOHYD 816 816 POTENTIAL.
FT CARBOHYD 842 842 POTENTIAL.
FT CARBOHYD 857 857 POTENTIAL.
SQ SEQUENCE 903 AA; 100305 MW; 65E6EC1C CRC32;

Query Match 81.2%; Score 173; DB 1; Length 903;
Best Local Similarity 74.2%; Pred. No. 5.16e-20;
Matches 23; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 651 DNGAGADTVKNDGVSRYFTDTRGNGRYSLK 681
QY 1 DNGAGADATKDDGVSRYFTTYDTNGRYSVK 31

RESULT 2
ID VGL_BPFI STANDARD; PRT: 348 AA.
AC P03657;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GENE I PROTEIN.
GN I.
OS BACTERIOPHAGE FL.
OC VIRUSES; SSDNA VIRUSES; INOVIIRIDAE; INOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 82211801.
RA BECK E., ZINK B.;
RT "Nucleotide sequence and genome organisation of filamentous
bacteriophages fl and fd.";
RL GENE 16:35-58(1981).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 83059882.
RA HILL D.F., PETERSEN G.B.;
RT "Nucleotide sequence of bacteriophage fl DNA.";
RL J. VIROL. 44:32-46(1982).
CC -!- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN ALTHOUGH
IT MAY BE INVOLVED IN PHASE ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00606; G14982; -
DR EMBL; J02448; G166210; -
DR PIR; C04262; Z1BPFI
KW PHAGE MATURATION; ATP-BINDING.
FT NP_BIND 8 15 ATP (POTENTIAL).
SQ SEQUENCE 348 AA; 39527 MW; 74E0133F CRC32;

Query Match 37.6%; Score 80; DB 1; Length 348;
Best Local Similarity 61.1%; Pred. No. 8.05e-02;
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 215 DSGVYS-YLTPYLSHGRY 231
QY 11 DGVYSRYFTTYDTNGRY 28

RESULT 3
ID VGL_BPM13 STANDARD; PRT: 348 AA.
AC P03656;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GENE I PROTEIN.
GN I.
OS BACTERIOPHAGE FL.
OC VIRUSES; SSDNA VIRUSES; INOVIIRIDAE; INOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 82211801.
RA BECK E., ZINK B.;
RT "Nucleotide sequence of bacteriophage fd DNA.";
RL NUCLEIC ACIDS RES. 5:4495-4503(1978).
CC -!- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN ALTHOUGH
IT MAY BE INVOLVED IN PHASE ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00602; G14940; -
DR EMBL; J02451; G215403; -
DR PIR; A04262; Z1BPFD.
KW PHAGE MATURATION; ATP-BINDING.
FT NP_BIND 8 15 ATP (POTENTIAL).
SQ SEQUENCE 348 AA; 39536 MW; 8400829E CRC32;

DE DE GENE I PROTEIN.
GN I.
OS BACTERIOPHAGE M13.
OC VIRUSES; SSDNA VIRUSES; INOVIIRIDAE; INOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 81067903.
RA VAN WEZENBEK P.M.G.F., HULSEBOS T.J.M., SCHOENMAKERS J.G.G.;
RT "Nucleotide sequence of the filamentous bacteriophage M13 DNA genome:
comparison with phage fd.";
RL GENE 11:129-148(1980).
CC -!- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN ALTHOUGH
IT MAY BE INVOLVED IN PHASE ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00604; G14968; -
DR PIR; B04262; Z1BPM3.
KW PHAGE MATURATION; ATP-BINDING.
FT NP_BIND 8 15 ATP (POTENTIAL).
SQ SEQUENCE 348 AA; 39550 MW; 2C14A910 CRC32;

Query Match 37.6%; Score 80; DB 1; Length 348;
Best Local Similarity 61.1%; Pred. No. 8.05e-02;
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 215 DSGVYS-YLTPYLSHGRY 231
QY 11 DGVYSRYFTTYDTNGRY 28

RESULT 4
ID VGL_BPFD STANDARD; PRT: 348 AA.
AC P03655;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE GENE I PROTEIN.
GN I.
OS BACTERIOPHAGE FD.
OC VIRUSES; SSDNA VIRUSES; INOVIIRIDAE; INOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=478 / HEIDELBERG;
RX MEDLINE; 79136480.
RA BECK E., SOMMER R., AUERSWALD E.A., KURZ C., ZINK B., OSTERBURG G.,
RA SCHALLER H., SUGIMOTO K., SUGISAKI H., OKAMOTO T., TAKANAMI M.;
RT "Nucleotide sequence of bacteriophage fd DNA.";
RL NUCLEIC ACIDS RES. 5:4495-4503(1978).
CC -!- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN ALTHOUGH
IT MAY BE INVOLVED IN PHASE ASSEMBLY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00602; G14940; -
DR EMBL; J02451; G215403; -
DR PIR; A04262; Z1BPFD.
KW PHAGE MATURATION; ATP-BINDING.
FT NP_BIND 8 15 ATP (POTENTIAL).
SQ SEQUENCE 348 AA; 39536 MW; 8400829E CRC32;
```



Query Match 37.6%; Score 80; DB 1; Length 348;  
Best Local Similarity 61.1%; Pred. No. 8.05e-02;  
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 215 DSGVYS-YLTPYLSHGRY 231  
| | | | | : | : | | |  
QY 11 DDGVYSRYFTYDTNGRY 28

RESULT 5  
ID CSG.METTE STANDARD; PRT; 593 AA.  
AC P27373;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE CELL SURFACE GLYCOPROTEIN PRECURSOR (S-LAYER PROTEIN).  
GN SLGA.  
OS METHANOTHERMUS FERVIDUS.  
OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOTHERMACEAE;  
OC METHANOTHERMUS.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-42.  
RC STRAIN=DSM 2088 / V245;  
RX BROECKL G., BEHR M., FABRY S., HENSEL R., KAUEWITZ H., BIENDL E.,  
RA KOENIG H.;  
RT "Analysis and nucleotide sequence of the genes encoding the surface-layer glycoproteins of the hyperthermophilic methanogens Methanothermus fervidus and Methanothermus sociabilis.";  
RL EUR. J. BIOCHEM. 199;147-152(1991).  
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.  
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIA IS COVERED BY A S-LAYER WITH HEXAGONAL SYMMETRY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X58297; G809714; -  
DR PIR: S16225; S16225.  
KW GLYCOPROTEIN; CELL WALL; S-LAYER; SIGNAL.  
FT SIGNAL 1 22  
FT CHAIN 23 593 CELL SURFACE GLYCOPROTEIN.  
FT CARBOHYD 29 29  
FT CARBOHYD 58 58  
FT CARBOHYD 66 66  
FT CARBOHYD 74 74  
FT CARBOHYD 114 114  
FT CARBOHYD 122 122  
FT CARBOHYD 145 145  
FT CARBOHYD 148 148  
FT CARBOHYD 158 158  
FT CARBOHYD 176 176  
FT CARBOHYD 208 208  
FT CARBOHYD 231 231  
FT CARBOHYD 326 326  
FT CARBOHYD 336 336  
FT CARBOHYD 340 340  
FT CARBOHYD 431 431  
FT CARBOHYD 471 471  
FT CARBOHYD 500 500  
FT CARBOHYD 516 516  
SQ SEQUENCE 593 AA; 65481 MW; 5CFA9AA9 CRC32;

Query Match 34.7%; Score 74; DB 1; Length 593;  
Best Local Similarity 35.3%; Pred. No. 7.32e-01;  
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 295 ASRNKIYGNFYNYET 311  
| : : : : | : | | | |  
QY 8 ATKDDGVYSRYFTTYDT 24

RESULT 6  
ID CSG.METSC STANDARD; PRT; 593 AA.  
AC P27374;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE CELL SURFACE GLYCOPROTEIN PRECURSOR (S-LAYER PROTEIN).  
GN SLGA.  
OS METHANOTHERMUS SOCIABILIS.  
OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOTHERMACEAE;  
OC METHANOTHERMUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 3496 / KF1-FL;  
RX MEDLINE: 91293115.  
RA BROECKL G., BEHR M., FABRY S., HENSEL R., KAUEWITZ H., BIENDL E.,  
RA KOENIG H.;  
RT "Analysis and nucleotide sequence of the genes encoding the surface-layer glycoproteins of the hyperthermophilic methanogens Methanothermus fervidus and Methanothermus sociabilis.";  
RL EUR. J. BIOCHEM. 199;147-152(1991).  
CC -!- SUBUNIT: ASSEMBLE INTO MONO-LAYERED CRYSTALLINE ARRAYS.  
CC -!- SUBCELLULAR LOCATION: CELL WALL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X58296; G809717; -  
DR PIR: S16375; S16375.  
KW GLYCOPROTEIN; CELL WALL; S-LAYER; SIGNAL.  
FT SIGNAL 1 22  
FT CHAIN 23 593 CELL SURFACE GLYCOPROTEIN.  
FT CARBOHYD 29 29  
FT CARBOHYD 58 58  
FT CARBOHYD 66 66  
FT CARBOHYD 74 74  
FT CARBOHYD 114 114  
FT CARBOHYD 122 122  
FT CARBOHYD 145 145  
FT CARBOHYD 148 148  
FT CARBOHYD 158 158  
FT CARBOHYD 176 176  
FT CARBOHYD 208 208  
FT CARBOHYD 231 231  
FT CARBOHYD 326 326  
FT CARBOHYD 336 336  
FT CARBOHYD 340 340  
FT CARBOHYD 431 431  
FT CARBOHYD 471 471  
FT CARBOHYD 500 500  
FT CARBOHYD 516 516  
SQ SEQUENCE 593 AA; 65503 MW; 52B1B8C8 CRC32;

Query Match 34.7%; Score 74; DB 1; Length 593;  
Best Local Similarity 35.3%; Pred. No. 7.32e-01;  
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 295 ASRNKIYGNFYNYET 311  
| : : : : | : | | | |  
QY 8 ATKDDGVYSRYFTTYDT 24

RESULT 7

```
ID AR56_SCHPO STANDARD; PRT: 885 AA.
AC P31318;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ARG11 PROTEIN PRECURSOR [CONTAINS: N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
DE REDUCTASE (EC 1.2.1.38) (N-ACETYL-GLUTAMATE SEMIALDEHYDE
DE DEHYDROGENASE) (NAGSA DEHYDROGENASE); ACETYLGLUTAMATE KINASE
DE (EC 2.7.2.8) (NAG KINASE) (AGK) (N-ACETYL-L-GLUTAMATE 5-
DE PHOSPHOTRANSFERASE)].
GN ARG11 OR SPAC49.09C
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC SCHIZOSACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=975;
RX MEDLINE; 92209520.
RA VAN HUFFEL C., DUBOIS E., MESSENGUY F.;
RT "Cloning and sequencing of arg3 and arg11 genes of
RT Schizosaccharomyces pombe on a 10-kb DNA fragment. Heterologous
RT expression and mitochondrial targeting of their translation
RT products".
RL EUR. J. BIOCHEM. 205:33-43(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED, (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- CATALYTIC ACTIVITY: N-ACETYL-L-GLUTAMATE 5-SEMIALDEHYDE + NADP(+)
CC + ORTHOPHOSPHATE = N-ACETYL-5-GLUTAMYL PHOSPHATE + NADPH.
CC -!- CATALYTIC ACTIVITY: ATP + N-ACETYL-L-GLUTAMATE = ADP +
CC N-ACETYL-L-GLUTAMATE 5-PHOSPHATE.
CC -!- ENZYME REGULATION: THE KINASE ACTIVITY IS INHIBITED BY ARGinine.
CC -!- PATHWAY: SECOND AND THIRD STEPS IN ARGinine BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- PTM: THE PROTEIN PRECURSOR IS CLEAVED INTO THE TWO BIOLOGICALLY
CC ACTIVE ENZYMES, THE KINASE AND THE REDUCTASE.
CC -!- SIMILARITY: WITH BACTERIAL ARG IN THE KINASE DOMAIN AND WITH
CC BACTERIAL ARG IN THE REDUCTASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X63576; G4906; -
CC DR EMBL; Z69727; E223735; -
CC DR PIR; S22389; S22389.
CC DR PFAM; PF01118; Semialdehyde_dh; 1.
CC DR OXIDOREDUCTASE; TRANSFERASE; KINASE; ARGinine BIOSYNTHESIS;
CC KW MITOCHONDRIUM; MULTIFUNCTIONAL ENZYME; TRANSIT PEPTIDE.
CC TRANSIT 1 59 MITOCHONDRIUM (POTENTIAL).
CC FT CHAIN 60 550 ACETYLGLUTAMATE KINASE (POTENTIAL).
CC FT CHAIN 551 885 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
CC REDUCTASE (POTENTIAL).
CC FT ACT_SITE 703 703 BY SIMILARITY.
CC FT VARIANT 261 261 V -> D (IN STRAIN 975).
CC FT VARIANT 345 345 G -> P (IN STRAIN 975).
CC SQ SEQUENCE 885 AA; 97704 MW; D5CD8C4C CRC32;

Query Match 34.7%; Score 74; DB 1; Length 885;
Best Local Similarity 22.2%; Pred. No. 7.32e-01;
Matches 6; Conservative 12; Mismatches 8; Indels 1; Gaps 1;

Db 70 SSIGSRREVQYLYRYFTSFPAQ-RFAI 95
: : : : | | | | : : : :
QY 4 AGADATKDGQVYSRYFTTYDTNGRYSV 30
```

```
RESULT 8
ID HIS8_ZYMMO STANDARD; PRT: 369 AA.
AC P34037;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-
DE PHOSPHATE TRANSAMINASE).
GN HIS8 OR HIS.
OS ZYMONOMAS MOBILIS.
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ZYMONOMAS GROUP;
OC ZYMONOMAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 31821 / CP4;
RX MEDLINE; 95189736.
RA GU W., ZHAO G.S., EDDY C., JENSEN R.A.;
RT "Imidazole acetol phosphate aminotransferase in Zymomonas mobilis:
RT molecular genetic, biochemical, and evolutionary analyses.".
RL J. BACTERIOL. 177:1576-1584(1995).
RN [2]
RP SEQUENCE OF 192-369 FROM N.A.
RC STRAIN-ATCC 31821 / CP4;
RX MEDLINE; 93183620.
RA ZHAO G., XIA T., INGRAM L.O., JENSEN R.A.;
RT "An allosterically insensitive class of cyclohexadienyl dehydrogenase
RT from Zymomonas mobilis.".
RL EUR. J. BIOCHEM. 212:157-165(1993).
CC -!- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGLUTARATE =
CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36343; G762852; -
CC DR EMBL; X67208; -; NOT_ANNOTATED_CDS.
CC DR PIR; S29383; S29383.
CC DR PROSITE; PS00599; AA_TRANSFERRER_CLASS_2; FALSE_NEG.
CC DR PFAM; PF00222; aminotran_2; 1.
CC DR HISTIDINE BIOSYNTHESIS; TRANSFERASE; AMINOTRANSFERASE;
CC KW PYRIDOXAL PHOSPHATE.
CC FT BINDING 280 280 PYRIDOXAL PHOSPHATE (PROBABLE).
CC SQ SEQUENCE 369 AA; 40177 MW; 3FCDAAB CRC32;

Query Match 34.3%; Score 73; DB 1; Length 369;
Best Local Similarity 48.1%; Pred. No. 1.05e+00;
Matches 13; Conservative 4; Mismatches 8; Indels 2; Gaps 2;

Db 98 AAGAGAGQDDEVLYPRYSFVSYPYLAAR 124
: | | | | | | | | | | | | | | : |
QY 3 GAGADATKDGQV-YSRY-FTTYDTNGR 27

RESULT 9
ID CH15_DROSU STANDARD; PRT: 121 AA.
AC P13426;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHORION PROTEIN S15.
GN CP15 OR S15.
OS DROSOPHILA SUBOBSCURA (FRUIT FLY).
```

```
[4] SEQUENCE FROM N.A.
RN RP STRAIN=IL125-2B;
RC MEDLINE; 91353083.
RA CHEN W., BALZI E., CAPIEAUX E., CHODER M., GOFFEAU A.;
RT "The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1
and ATE1 loci on chromosome VII from Saccharomyces cerevisiae reveals
the PDR6 gene, a new member of the genetic network controlling
pleiotropic drug resistance.";
RT YEAST 7:287-299(1991).
RL [5]
RN SEQUENCE FROM N.A.
RP HEBLING U., HOFMANN B., DELIUS H.;
RA SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RL [6]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 10-252.
RP GROLL M., DITZEL L., LOWE J., STOCK D., BOCHTLER M., BARTUNIK H.D.,
RA MEDLINE; 97242404.
RT "Structure of 20S proteasome from yeast at 2.4-A resolution.";
RN NATURE 386:463-471(1997).
CC -!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH
ARG. PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT
NEUTRAL OR SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT
PROTEOLYTIC ACTIVITY.
CC -!- PATHWAY: IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL
PROTEOLYTIC PATHWAY.
CC -!- SUBUNIT: YEAST PROTEASOME SEEMS TO BE COMPOSED OF 14 DIFFERENT
SUBUNITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY T1A; ALSO KNOWN AS THE
PROTEASOME A-TYPE FAMILY. BELONGS TO THE IOTA SUBFAMILY.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN IN THE N-TERMINUS
DUE TO A FRAMESHIFT.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M55440; G173206; -
DR EMBL; M63641; G172266; -
DR EMBL; X56732; G506484; -
DR EMBL; M31430; G172546; ALT_FRAME.
DR EMBL; S58126; E9136; -
DR EMBL; Z72533; E243317; -
DR PIR; S11199; SNBYS1.
DR PIR; S31557; S31557.
DR PDB; 1RYP; 15-APR-98.
DR SGD; L0001812; SCLI.
DR PROSITE; PS00388; PROTEASOME_A; 1.
DR PFAM; PF00227; Proteasome; 1.
DR PROTEASOME; HYDROLASE; PROTEASE; 3D-STRUCTURE.
SQ SEQUENCE 252 AA; 28001 MW; 5748C824 CRC32;

Query Match      33.3%; Score 71; DB 1; Length 252;
Best Local Similarity 39.3%; Pred. No. 2.12e+00;
Matches 11; Conservative 5; Mismatches 10; Indels 2; Gaps 2;

Db       2 SCNAASAAG-YDRHITIFSPGRLYQV 28
        :|||: | |||: | |: ||| | |
QY       4 AGADATKDDGVSYRYFTTYDTNGR-YSV 30

RESULT 11
ID YHX2_YEAST STANDARD; PRT; 823 AA.
AC P38863;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
```

```

01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
HYPOTHETICAL 96.8 KD PROTEIN IN NMD3-ENO2 INTERGENIC REGION.
YHR172W.
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCHAROMYCETALES;
SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
MEDLINE: 94378003.
JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
DU Z., FAVELLO A., FULLTON L., GATTUNG S., GEISEL C., KRISTEN J.,
KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASCHI E., VAUGHAN K.,
VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
VAUDIN M.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
SCIENCE 265:2077-2082(1994).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC      EMBL: U00027; G458898; -
DR      PIR: S48911; S48911.
KW      HYPOTHETICAL PROTEIN.
FT      HYDROPHOBIC REGION.
SQ      SEQUENCE 823 AA; 96824 MW; 34DF9955 CRC32;

Query Match          32.9%; Score 70; DB 1; Length 823;
Best Local Similarity 42.9%; Pred. No. 3.00e+00;
Matches           6; Conservative         3; Mismatches   0; Indels   0; Gaps   0;

Db    71 ECTYRINDPEPS 84
QY    12 DGYSRYFTYDTN 25

RESULT 12
ID     SAOX.ECOLI STANDARD; PRT; 372 AA.
AC     P40874; Q47144;
DT     01-NOV-1997 (REL. 35, CREATED)
DT     01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT     01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE     PUTATIVE SARCOSE OXIDASE (EC 1.5.3.1).
GN     SOLA.
OS     ESCHERICHIA COLI.
OC     BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC     ESCHERICHIA.
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=K12 / MG1655;
RC     MEDLINE: 97426617.
RX     BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA     RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA     GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA     MAU B., SHAO Y.;
RT     "The complete genome sequence of Escherichia coli K-12."
RL     SCIENCE 277:1453-1474(1997).
CC     -!- CATALYTIC ACTIVITY: SARCOSE + H(2)O + O(2) = GLYCINE +
CC     FORMALDEHYDE + H(2)O(2).
CC     -!- COFACTOR: FAD FLAVOPROTEIN (1 MOLE OF FAD PER MOLE OF ENZYME).
CC     -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC     -----

01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
HYPOTHETICAL 96.8 KD PROTEIN IN NMD3-ENO2 INTERGENIC REGION.
YHR172W.
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCHAROMYCETALES;
SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
MEDLINE: 94378003.
JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
DU Z., FAVELLO A., FULLTON L., GATTUNG S., GEISEL C., KRISTEN J.,
KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASCHI E., VAUGHAN K.,
VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
VAUDIN M.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
SCIENCE 265:2077-2082(1994).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC      EMBL: U00027; G458898; -
DR      PIR: S48911; S48911.
KW      HYPOTHETICAL PROTEIN.
FT      HYDROPHOBIC REGION.
SQ      SEQUENCE 823 AA; 96824 MW; 34DF9955 CRC32;

Query Match          32.9%; Score 70; DB 1; Length 823;
Best Local Similarity 42.9%; Pred. No. 3.00e+00;
Matches           6; Conservative         3; Mismatches   0; Indels   0; Gaps   0;

Db    71 ECTYRINDPEPS 84
QY    12 DGYSRYFTYDTN 25

RESULT 12
ID     SAOX.ECOLI STANDARD; PRT; 372 AA.
AC     P40874; Q47144;
DT     01-NOV-1997 (REL. 35, CREATED)
DT     01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT     01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE     PUTATIVE SARCOSE OXIDASE (EC 1.5.3.1).
GN     SOLA.
OS     ESCHERICHIA COLI.
OC     BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC     ESCHERICHIA.
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=K12 / MG1655;
RC     MEDLINE: 97426617.
RX     BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA     RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA     GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA     MAU B., SHAO Y.;
RT     "The complete genome sequence of Escherichia coli K-12."
RL     SCIENCE 277:1453-1474(1997).
CC     -!- CATALYTIC ACTIVITY: SARCOSE + H(2)O + O(2) = GLYCINE +
CC     FORMALDEHYDE + H(2)O(2).
CC     -!- COFACTOR: FAD FLAVOPROTEIN (1 MOLE OF FAD PER MOLE OF ENZYME).
CC     -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC     -----

01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
HYPOTHETICAL 96.8 KD PROTEIN IN NMD3-ENO2 INTERGENIC REGION.
YHR172W.
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCHAROMYCETALES;
SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
MEDLINE: 94378003.
JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
DU Z., FAVELLO A., FULLTON L., GATTUNG S., GEISEL C., KRISTEN J.,
KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASCHI E., VAUGHAN K.,
VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
VAUDIN M.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
SCIENCE 265:2077-2082(1994).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC      EMBL: U00027; G458898; -
DR      PIR: S48911; S48911.
KW      HYPOTHETICAL PROTEIN.
FT      HYDROPHOBIC REGION.
SQ      SEQUENCE 823 AA; 96824 MW; 34DF9955 CRC32;

Query Match          32.9%; Score 70; DB 1; Length 823;
Best Local Similarity 42.9%; Pred. No. 3.00e+00;
Matches           6; Conservative         3; Mismatches   0; Indels   0; Gaps   0;

Db    71 ECTYRINDPEPS 84
QY    12 DGYSRYFTYDTN 25

RESULT 12
ID     SAOX.ECOLI STANDARD; PRT; 372 AA.
AC     P40874; Q47144;
DT     01-NOV-1997 (REL. 35, CREATED)
DT     01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT     01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE     PUTATIVE SARCOSE OXIDASE (EC 1.5.3.1).
GN     SOLA.
OS     ESCHERICHIA COLI.
OC     BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC     ESCHERICHIA.
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=K12 / MG1655;
RC     MEDLINE: 97426617.
RX     BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA     RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA     GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA     MAU B., SHAO Y.;
RT     "The complete genome sequence of Escherichia coli K-12."
RL     SCIENCE 277:1453-1474(1997).
CC     -!- CATALYTIC ACTIVITY: SARCOSE + H(2)O + O(2) = GLYCINE +
CC     FORMALDEHYDE + H(2)O(2).
CC     -!- COFACTOR: FAD FLAVOPROTEIN (1 MOLE OF FAD PER MOLE OF ENZYME).
CC     -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC     -----

01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
HYPOTHETICAL 96.8 KD PROTEIN IN NMD3-ENO2 INTERGENIC REGION.
YHR172W.
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACCHAROMYCETALES;
SACCHAROMYCETACEAE; SACCHAROMYCES.
[1]
SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
MEDLINE: 94378003.
JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,
DU Z., FAVELLO A., FULLTON L., GATTUNG S., GEISEL C., KRISTEN J.,
KUCABA T., HILLIER L., JIER M., JOHNSTON L., LANGSTON Y.,
LATREILLE P., LOUIS E.J., MACRI C., MARDIS E., MENEZES S., MOUSER L.,
NHAN M., RIFKIN L., RILES L., ST PETER H., TREVASCHI E., VAUGHAN K.,
VIGNATI D., WILCOX L., WOHLDMAN P., WATERSTON R., WILSON R.,
VAUDIN M.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
VIII."
SCIENCE 265:2077-2082(1994).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC      E
```

DR PFAM; PF00001: 7tm.1; 1.  
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.  
FT EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 1 39  
FT TRANSSEM 40 67  
FT DOMAIN 68 77  
FT TRANSSEM 78 98  
FT DOMAIN 99 111  
FT TRANSSEM 112 133  
FT DOMAIN 134 150  
FT TRANSSEM 151 175  
FT DOMAIN 176 206  
FT TRANSSEM 207 226  
FT DOMAIN 227 242  
FT TRANSSEM 243 267  
FT DOMAIN 268 284  
FT TRANSSEM 285 308  
FT DOMAIN 309 360  
FT CARBOHYD 183 183  
FT CARBOHYD 194 194  
FT DISULFID 110 187  
SQ SEQUENCE 360 AA; 41402 MW; 8738E75E CRC32;

Query Match 31.9%; Score 68; DB 1; Length 360;  
Best Local Similarity 42.1%; Pred. No. 5.96e+00;  
Matches 8; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Db 7 ADTLDSEISNYL-YES 24  
QY 6 ADATKDDGVSRVFTTYDT 24

RESULT 14  
ID VP2\_HPV22 STANDARD; PRT; 436 AA.  
AC P50768;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE REGULATORY PROTEIN E2.  
GN HUMAN PAPILLOMAVIRUS TYPE 22.  
OS VIRUSES; DSDNA VIRUSES, NO RNA STAGE; PAPOVAVIRIDAE; PAPILLOMAVIRUS.  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; PAPOVAVIRIDAE; PAPILLOMAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA DELIUS H.;  
RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U31780; G1020182; -  
DR PFAM; PF00508; E2\_N; 1.  
DR PFAM; PF00511; E2\_C; 1.  
DR HSSP; P03122; 2BOP.  
SQ SEQUENCE 436 AA; 49728 MW; E6906B6A CRC32;

Query Match 31.5%; Score 67; DB 1; Length 436;  
Best Local Similarity 26.3%; Pred. No. 8.35e+00;  
Matches 5; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

Db 162 EGTFTKTYIKFETDAKRYG 180  
QY 12 DGVYSRFTTYDING-RYS 29

RESULT 15  
ID Y944\_HELPY STANDARD; PRT; 125 AA.  
AC O25598;  
DT 15-DEC-1998 (REL. 37, CREATED)

DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL PROTEIN HP0944.  
GN HP0944.  
OS BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;  
OC HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).  
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;  
OC HELICOBACTER.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE; 97394467.  
RA FLEISCHMANN R.D., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
RA TOMB J.-F., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,  
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,  
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,  
RA VENTER J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori";  
RL NATURE 388:539-547(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE000603; G2314082; -  
DR TIGR; HP0944; -  
DR PFAM; PF01042; DUF10; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 125 AA; 13373 MW; D2347CCC CRC32;  
Query Match 31.0%; Score 66; DB 1; Length 125;  
Best Local Similarity 35.0%; Pred. No. 1.17e+01;  
Matches 7; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Db 89 NGIYGSVFTEPYPARATFOV 108  
QY 12 DGVYSRIFT-TYDINGRYSV 30

Search completed: Mon Aug 23 13:15:13 1999  
Job time : 11 secs.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:15:32 1999; MasPar time 6.90 Seconds

Tabular output not generated. 245.138 Million cell updates/sec

Title: >US-09-049-696-47  
Perfect Score: 213  
Sequence: 1 DNGAGADATKDDGVYSRYFTYDTNGRYSVK 31

Scoring table: PAM 150  
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spbrembl9  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 31.105; Variance 55.696; scale 0.558

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	194	91.1	913	11	088826	6.43e-23
2	169	79.3	901	11	088860	7.98e-18
3	164	77.0	794	6	018742	8.01e-17
4	164	77.0	820	6	018743	8.01e-17
5	164	77.0	905	6	018741	8.01e-17
6	75	35.2	137	2	050960	1.12e+00
7	75	35.2	245	1	028187	1.12e+00
8	75	35.2	353	9	080239	1.12e+00
9	74	34.7	337	2	034232	1.58e+00
10	74	34.7	337	2	087136	1.58e+00
11	72	33.8	2958	3	012150	3.14e+00
12	71	33.3	57	10	038932	4.40e+00
13	70	32.9	120	1	028622	6.15e+00
14	70	32.9	245	9	038204	6.15e+00
15	70	32.9	328	11	088775	6.15e+00
16	70	32.9	603	2	070041	6.15e+00
17	70	32.9	731	2	070039	6.15e+00
18	70	32.9	1274	10	P93125	6.15e+00
19	69	32.4	336	2	050359	8.57e+00
20	69	32.4	350	14	082008	8.57e+00

21 69 32.4 365 2 085800 RESTRICTION-MODIFICATI 8.57e+00  
22 69 32.4 366 2 085797 RESTRICTION-MODIFICATI 8.57e+00  
23 69 32.4 369 2 030382 SPECIFICITY (S) SUBUNI 8.57e+00  
24 69 32.4 398 2 085798 RESTRICTION-MODIFICATI 8.57e+00  
25 69 32.4 399 2 085799 RESTRICTION-MODIFICATI 8.57e+00  
26 69 32.4 401 2 050356 RESTRICTION-MODIFICATI 8.57e+00  
27 69 32.4 572 3 087247 RHAMNOLACTURONAN HYD 8.57e+00  
28 69 32.4 1577 2 055265 GLUCOSYLTRANSFERASE PR 8.57e+00  
29 68 31.9 169 2 055969 HYPOTHETICAL 19.2 KD P 1.19e+01  
30 68 31.9 309 2 085832 HYPOTHETICAL 36.1 KD P 1.19e+01  
31 68 31.9 360 14 072710 PAKISTAN STRAIN COMPLE 1.19e+01  
32 68 31.9 360 14 072707 PAKISTAN STRAIN COMPLE 1.19e+01  
33 68 31.9 546 9 080243 MARRP. 1.19e+01  
34 68 31.9 807 5 077341 MAL3P4.5 PROTEIN. 1.19e+01  
35 68 31.9 1214 5 025338 DELTA-LATROINSECTOTOXI 1.19e+01  
36 67 31.5 150 14 067622 ORF 1742. 1.65e+01  
37 67 31.5 403 1 026937 PROTEIN-EXPORT MEMBRAN 1.65e+01  
38 67 31.5 478 5 062571 STRESS-RESPONSIVE PROT 1.65e+01  
39 67 31.5 479 5 090832 F10C2.7 PROTEIN. 1.65e+01  
40 67 31.5 603 4 060843 PUTATIVE SERINE/THREON 1.65e+01  
41 67 31.5 716 3 013477 PUTATIVE GLUCOSE SENSO 1.65e+01  
42 67 31.5 1099 2 054377 EXONUCLEASE REXB. 1.65e+01  
43 67 31.5 1289 5 017174 GP80 (FRAGMENT). 1.65e+01  
44 66 31.0 632 10 082261 F17A22.33 PROTEIN. 2.28e+01  
45 66 31.0 1704 2 051816 ARG-GINGIPAIN-1 PROTEI 2.28e+01

## ALIGNMENTS

RESULT 1  
ID 088826 PRELIMINARY; PRT; 913 AA.  
AC 088826;  
DT 01-NOV-1998 (TREMREL. 08, CREATED)  
DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE GOB-5 PROTEIN.  
GN GOB-5.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RA KOMIYA T., TANIGAWA Y., HIROHASHI S.;  
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet cells in mice."  
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AB017156; D1034712;  
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 91.1%; Score 194; DB 11; Length 913;  
Best Local Similarity 83.9%; Pred. No. 6.43e-23;  
Matches 26; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 645 DNGAGADATKDDGVYSRYFTYDTNGRYSVK 675  
QY 1 DNGAGADATKDDGVYSRYFTYDTNGRYSVK 31  
|||||

RESULT 2  
ID 088860 PRELIMINARY; PRT; 901 AA.  
AC 088860;  
DT 01-NOV-1998 (TREMREL. 08, CREATED)  
DT 01-NOV-1998 (TREMREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE CHLORIDE CHANNEL CACC.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ROMIO L., MUSANTE L., CINTI R., MORAN O., SERI M., GALIETTA L.J.V.;  
RT "Characterization of a murine gene homologous to the bovine CaCC

RT chloride channel.";  
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF052746; G3560547; -.  
SQ SEQUENCE 901 AA; 100039 MW; E52BF02E CRC32;

Query Match 79.3%; Score 169; DB 11; Length 901;  
Best Local Similarity 67.7%; Pred. No. 7.98e-18;  
Matches 21; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 649 DNGAGDVTKNDGIYRYFTDYHGNSYSLK 679  
QY 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31

RESULT 3

ID O18742 PRELIMINARY; PRT: 794 AA.

AC O18742;

DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;  
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF001262; G2623765; -.  
SQ SEQUENCE 794 AA; 88509 MW; 2D450EB CRC32;

Query Match 77.0%; Score 164; DB 6; Length 794;  
Best Local Similarity 71.0%; Pred. No. 8.01e-17;  
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGDVTKNDGIYRYFTDYHGNSYSLK 682  
QY 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31

RESULT 4

ID O18743 PRELIMINARY; PRT: 820 AA.

AC O18743;

DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
DE LU-ECAM-1.

OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;  
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF001263; G2623767; -.  
SQ SEQUENCE 820 AA; 91464 MW; 171ED21C CRC32;

Query Match 77.0%; Score 164; DB 6; Length 820;  
Best Local Similarity 71.0%; Pred. No. 8.01e-17;  
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGDVTKNDGIYRYFTDYHGNSYSLK 682  
QY 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31

RESULT 5

ID O18741 PRELIMINARY; PRT: 905 AA.

AC O18741;

DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE LU-ECAM-1.  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,  
RA GOODWIN A., PAULI B.U.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: AF001251; G2623763; -.

DR PFAM: PF00041; fn3; 1.

SQ SEQUENCE 905 AA; 101005 MW; 86E65459 CRC32;

Query Match 77.0%; Score 164; DB 6; Length 905;  
Best Local Similarity 71.0%; Pred. No. 8.01e-17;  
Matches 22; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 652 DNGAGDVTKNDGIYRYFTDYHGNSYSLK 682  
QY 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31

RESULT 6

ID O50960 PRELIMINARY; PRT: 137 AA.

AC O50960;

DT 01-JUN-1998 (TREMREL. 06, CREATED)  
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 16.4 KD PROTEIN.

GN BBA71.

OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).

OC PLASMID LP54

OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35210 / B31;

RX MEDLINE: 98065943.

RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,  
RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,  
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D., HANSON M.,  
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S.,  
RA VAN VUUT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,  
RA UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,  
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,  
RA SMITH H.O., VENTER J.C.;

RT "Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi."

RL NATURE 390:580-586(1997).

DR EMBL: AE000790; G2690288; -.

DR TIGR: BBA71; -.

KW HYPOTHETICAL PROTEIN; PLASMID.

SQ SEQUENCE 137 AA; 16412 MW; 36EB8C93 CRC32;

Query Match 35.2%; Score 75; DB 2; Length 137;  
Best Local Similarity 28.6%; Pred. No. 1.12e+00;  
Matches 6; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Db 104 KNEDAYKKYIAYINENYQI 124  
QY 10 KDDGVYSRYFTTYDNGRYSV 30

RESULT 7

ID O28187 PRELIMINARY; PRT: 245 AA.

AC O28187;

DT 01-JAN-1998 (TREMREL. 05, CREATED)  
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 28.6 KD PROTEIN.



```

GN AF2093.
OS ARCHAEOGLOBUS FULGIDUS.
OC ARCHAEA; EURYARCHAEOTA; ARCHAEOGLOBALES; ARCHAEOGLOBACEAE;
OC ARCHAEOGLOBUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RC MEDLINE: 98049343.
RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
RA FLEISCHMANN R.D., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
RA KIRKNESS E.F., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
RA PETERSON S., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
RA OYERBEEK R., SPRIGGS T., ARTIACH P., KATSEV S.M.,
RA COTTON M.D., SADOW P.W., D'ANDREA K.P., BOWMAN C., GARLAND S.A.,
RA SADON P.W., D'ANDREA K.P., BOWMAN C., GARLAND S.A.,
RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
RA VENTER J.C.;
RA "The complete genome sequence of the hyperthermophilic,
RT sulphate-reducing archaeon Archaeoglobus fulgidus."
RL NATURE 390:364-370(1997).
RL EMBL: AE000959; G264844;
DR TIGR: AF2093;
DR HYPOTHETICAL PROTEIN.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 245 AA; 28593 MW; 6093807B CRC32;

Query Match 35.2%; Score 75; DB 1; Length 245;
Best Local Similarity 46.7%; Pred. No. 1.12e+00;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 114 VDRFFEGYEAERY 128
QY 14 VYSRYFTYDNGRY 28

RESULT 8
ID U80299 PRELIMINARY; PRT; 353 AA.
AC U80299;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MATURATION PROTEIN.
DE GENE 1.
GN BACTERIOPHAGE IF1.
OS VIRUSES; SS DNA VIRUSES; INOVIRIDAE; INOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA HILL D.F., HUGHES G., MCNAUGHTON J.C., STOCKWELL P.A., PETERSEN G.B.;
RT "A sequence of the filamentous coliphage If1."
RL SUBMITTED (OCT-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U02303; G367628;
SQ SEQUENCE 353 AA; 39990 MW; 898679C9 CRC32;

Query Match 35.2%; Score 75; DB 9; Length 353;
Best Local Similarity 42.9%; Pred. No. 1.12e+00;
Matches 9; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

Db 215 EHSFYS-YLTPYLSHGKRYAVK 234
QY 11 DDGVYSRYFTYDNGRYSVK 31

RESULT 9
ID Q34232 PRELIMINARY; PRT; 337 AA.
AC Q34232;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF40X1 PROTEIN.
OS VIRUSIO CHOLERAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; VIBRIONACEAE; VIBRIO.
RN [1]

```

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

ID 088775

ID 088775

AC 088775;  
DT 01-NOV-1998 (TREMBLEL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLEL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLEL. 08, LAST ANNOTATION UPDATE)  
DE EMBIGIN PROTEIN PRECURSOR.  
GN EMBIGIN.  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE DAWLEY; TISSUE=REGRESSING PROSTATE;  
RA GUENETTE R.;  
RL SUBMITTED (JUL-1998) TO ENBL/GENEBANK/DBBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE DAWLEY; TISSUE=REGRESSING PROSTATE;  
RX MEDLINE: 98101284.  
RA GUENETTE R.; SRIDHAR S., HERLEY M., MOOIBROEK M., WONG P.,  
RA TENNISWOOD M.;  
RT "Embigin, A Developmentally Expressed Member of the Immunoglobulin  
RT Super Family, Is also Expressed During Regression of Prostate and  
RT Mammary Gland.";  
RL DEV. GENET. 21:268-278(1997).  
DR EMBL: AJ009698; E1312986; .  
KW SIGNAL.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 328 EMBIGIN PROTEIN.  
SQ SEQUENCE 328 AA; 37005 MW; 3BD13009 CRC32;  
  
Query Match 32.9%; Score 70; DB 11; Length 328;  
Best Local Similarity 36.4%; Pred. No. 6.15e+00;  
Matches 12; Conservative 10; Mismatches 7; Indels 4; Gaps 3;  
  
Db 111 ETTDGFNTKMGDTLYSQYRFTVNSKMGKYS 143  
: |::||| |::||| |::||| |::|||  
QY 1 DNGAGADATK-DDGVYSRY-FTTYDTN--GRYS 29

Search completed: Mon Aug 23 13:15:55 1999  
Job time : 23 secs.

***This Page Blank (uspto)***

\*\*\*\*\*

WQSEH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:03:56 1999; MasPar time 9.73 Seconds

Tabular output not generated. 67.739 Million cell updates/sec

Title: >US-09-049-696-47

Description: (1-31) from US09049696.pep

Perfect Score: 31  
Sequence: 1 DNGAGADATKDDGVYSRYFTTYDNGRYSVK 31

Scoring table: TABLE unitprotatable

Gap 60

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database: a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 1.991; Variance 0.666; scale 2.991

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	18	58.1	228	20	W06548 Human colon specific	2.66e-12
2	18	58.1	228	29	W46879 Protein sequence enco	2.66e-12
3	6	19.4	345	31	W41166 Metal-regulated trans	5.44e-01
4	6	19.4	372	28	W14284 N-methylamino acid ox	5.44e-01
5	6	19.4	480	21	W14499 Bacillus amyloliquefa	5.44e-01
6	6	19.4	480	14	R78268 Bacillus amyloliquefa	5.44e-01
7	6	19.4	483	28	W31405 Bacillus amyloliquefa	5.44e-01
8	6	19.4	483	14	R72448 Bacillus amyloliquefa	5.44e-01
9	6	19.4	514	2	R10378 Mutant alpha-amyliase	5.44e-01
10	6	19.4	514	2	R10379 Wild type alpha-amyliase	5.44e-01
11	6	19.4	520	28	W39743 B. amyloliquefaciens	5.44e-01
12	6	19.4	615	2	R10100 E. tenella protein Etl	5.44e-01
13	6	19.4	674	37	W72748 Human P-dlg protein	5.44e-01
14	6	19.4	729	17	R89275 Yeast coagulation pro	5.44e-01
15	6	19.4	996	28	W33624 Eimeria tenella 45 kD	5.44e-01
16	6	19.4	998	28	W33621 Eimeria tenella 45 kD	5.44e-01

Sequence of the antib  
Platenolide synthase  
Platenolide synthase  
Hybrid srmg/tylg ORF1  
Platenolide synthase  
Platenolide synthase  
T. niveum Cyclosporin  
Nephila clavipes spid  
JC virus VP-1 protein  
CD44 peptide CD44-1.  
Actin 1-28 peptide.  
Endoglycoceramidase f  
TSAR binding domain e  
Synthetic library pep  
Cystatin SN anticarie  
Secretory signal sequ  
Human salivary cystat  
Variable segment enco  
H. pylori cytoplasmic  
PT-NANB virus protein  
Amino acid sequence e  
Cytochrome c3 protein  
Desulfovibrio vulgari  
Variable region enco  
Cytochrome c3 protein  
Cytochrome c3 protein  
Protein sequence of d  
Vdelat (genomic).  
T-cell specific pepti  
Streptococcus pneumon  
Silk like protein (SL  
Consensus sequence of  
Streptococcus pneumon  
Borrelia SIMON anti  
Borrelia SIMON anti  
Sequence of novel chi  
Recombinant human glu  
Human glutathione per  
Human glutathione per  
Glutathione peroxidase  
Sequence of human glu  
H. pylori ORF 07ap110  
H. pylori protein.  
Deduced peptide sequ  
Minor ampullate spide  
H. pylori ORF 07ap110  
Modified tobacco etch  
TEV CP.  
Pamelyi CoA synthase.  
TrpA protein.  
Helicobacter pylori A  
Staphylococcus aureus  
Staphylococcus aureus  
Staphylococcus aureus  
M. tuberculosis immun  
Mycobacterium tubercu  
Staphylococcus aureus  
hTCP protein fragment  
H. pylori vacA protei  
TCR delta protein.  
Human islet cell anti  
Staphylococcus aureus  
Residual protease-2  
Residual protease II.  
Racillus subtilis RP-  
PEP PM.  
B. lactofermentum pre  
Feedback inhibition-r  
CM-PP gene product (S  
CM-PP gene product (S  
PAJ16 fragment.  
1-Aminocyclopropane-1  
H. pylori surface or  
Glutamine synthetase

998 4 P93706  
1936 26 W22607  
1996 26 W23717  
4545 26 W22611  
4550 26 W23716  
4550 26 W22606  
15281 9 R44929  
9 30 W53358  
20 24 W19594  
22 10 R53473  
28 24 W15768  
38 11 R58400  
42 18 R95538  
56 11 R60289  
69 2 P70041  
80 7 R37317  
89 1 P95053  
90 22 W20636  
94 7 R37858  
94 1 R05262  
107 30 W57220  
107 30 W57218  
107 1 P95051  
107 30 W57219  
107 30 W57221  
117 39 W86141  
120 6 R32911  
123 18 R95874  
130 33 W61016  
142 19 R95135  
153 10 R54634  
160 32 W58554  
178 12 R62776  
178 12 R60890  
185 1 P82003  
200 1 P80436  
201 20 W08076  
201 17 R89910  
201 13 R66235  
201 2 P80918  
211 29 W55324  
211 22 W20755  
232 1 P80760  
251 16 R80192  
257 29 W55462  
263 20 W12466  
263 8 R45612  
276 18 W00104  
280 19 W01142  
280 37 W72879  
284 39 W89750  
284 39 W89751  
284 37 W81733  
284 34 W64366  
287 31 W28097  
288 27 W34527  
288 28 W27720  
293 6 R32905  
298 26 W35300  
309 39 R89756  
313 1 R04905  
313 21 W01791  
313 23 W16338  
313 7 R37246  
315 21 W01677  
315 21 W01678  
315 9 R47927  
315 9 R47926  
315 2 R24019  
320 17 R91306  
328 22 W20269  
337 1 P91388

17 6 19.4  
18 6 19.4  
19 6 19.4  
20 6 19.4  
21 6 19.4  
22 6 19.4  
23 6 19.4  
24 5 16.1  
25 5 16.1  
26 5 16.1  
27 5 16.1  
28 5 16.1  
29 5 16.1  
30 5 16.1  
31 5 16.1  
32 5 16.1  
33 5 16.1  
34 5 16.1  
35 5 16.1  
36 5 16.1  
37 5 16.1  
38 5 16.1  
39 5 16.1  
40 5 16.1  
41 5 16.1  
42 5 16.1  
43 5 16.1  
44 5 16.1  
45 5 16.1  
46 5 16.1  
47 5 16.1  
48 5 16.1  
49 5 16.1  
50 5 16.1  
51 5 16.1  
52 5 16.1  
53 5 16.1  
54 5 16.1  
55 5 16.1  
56 5 16.1  
57 5 16.1  
58 5 16.1  
59 5 16.1  
60 5 16.1  
61 5 16.1  
62 5 16.1  
63 5 16.1  
64 5 16.1  
65 5 16.1  
66 5 16.1  
67 5 16.1  
68 5 16.1  
69 5 16.1  
70 5 16.1  
71 5 16.1  
72 5 16.1  
73 5 16.1  
74 5 16.1  
75 5 16.1  
76 5 16.1  
77 5 16.1  
78 5 16.1  
79 5 16.1  
80 5 16.1  
81 5 16.1  
82 5 16.1  
83 5 16.1  
84 5 16.1  
85 5 16.1  
86 5 16.1  
87 5 16.1  
88 5 16.1  
89 5 16.1

90	5	16.1	338 22	W21029	H. pylori transmembr	4.68e+02	163	5	16.1	483 38	W73510	Alpha-amylase protein	4.68e+02
91	5	16.1	348 29	W52169	Human MKK3-interactin	4.68e+02	164	5	16.1	483 6	R34289	B.licheniformis alpha	4.68e+02
92	5	16.1	354 35	W74521	Human ATP-sensitive p	4.68e+02	165	5	16.1	483 28	R31404	Bacillus licheniformi	4.68e+02
93	5	16.1	354 26	W24016	Human JC virus VP1 pr	4.68e+02	166	5	16.1	483 17	R88451	Alpha-amylase V128E/H	4.68e+02
94	5	16.1	354 27	W30915	JC virus VP1 protein.	4.68e+02	167	5	16.1	483 11	R57993	M138X alpha-amylase.	4.68e+02
95	5	16.1	356 2	W27164	Glutamine synthetase.	4.68e+02	168	5	16.1	483 14	R82267	Bacillus licheniformi	4.68e+02
96	5	16.1	356 5	P70469	Sequence of alfafa gl	4.68e+02	169	5	16.1	483 22	W05231	Bacillus licheniformi	4.68e+02
97	5	16.1	356 2	P70836	Sequence of alfafa gl	4.68e+02	170	5	16.1	483 18	R81479	Mutant mature alpha a	4.68e+02
98	5	16.1	356 2	P71066	Sequence of mutant gl	4.68e+02	171	5	16.1	483 6	R31363	B.licheniformis alpha	4.68e+02
99	5	16.1	356 2	R06491	Glutamine synthase in	4.68e+02	172	5	16.1	483 11	R58718	M197N alpha-amylase.	4.68e+02
100	5	16.1	357 2	R11509	Recombinant iso-penic	4.68e+02	173	5	16.1	483 11	R57999	M8A alpha-amylase.	4.68e+02
101	5	16.1	357 8	R44216	Streptomyces fradiae	4.68e+02	174	5	16.1	483 11	R57994	M197X alpha-amylase.	4.68e+02
102	5	16.1	357 2	R06492	Glutamine synthase in	4.68e+02	175	5	16.1	483 33	W57976	Alpha-amylase mutant	4.68e+02
103	5	16.1	358 1	P93169	Lettuce glutamine syn	4.68e+02	176	5	16.1	483 11	R58729	M15H alpha-amylase.	4.68e+02
104	5	16.1	361 18	R91444	Human haematopoietic	4.68e+02	177	5	16.1	483 11	R58722	M197C alpha-amylase.	4.68e+02
105	5	16.1	361 4	R20816	Haematopoietic CD44 A	4.68e+02	178	5	16.1	483 11	R58000	B.licheniformis alpha	4.68e+02
106	5	16.1	361 8	R42281	EGE2 receptor coding	4.68e+02	179	5	16.1	483 6	R34288	B.licheniformis alpha	4.68e+02
107	5	16.1	362 1	R07355	B7 adhesion receptor.	4.68e+02	180	5	16.1	483 11	R57995	M15X alpha-amylase.	4.68e+02
108	5	16.1	364 35	W74518	Human ATP-sensitive p	4.68e+02	181	5	16.1	483 11	R58013	M197H alpha-amylase.	4.68e+02
109	5	16.1	365 8	R42280	EGE2 receptor coding	4.68e+02	182	5	16.1	483 11	R58747	W138H alpha-amylase.	4.68e+02
110	5	16.1	366 32	W37945	Amino acid sequence o	4.68e+02	183	5	16.1	483 11	R58732	M15S alpha-amylase.	4.68e+02
111	5	16.1	370 39	W95181	Human G-protein coupl	4.68e+02	184	5	16.1	483 11	R58741	M15W alpha-amylase.	4.68e+02
112	5	16.1	370 23	W18093	Type I diabetes-assoc	4.68e+02	185	5	16.1	483 11	R58725	M15Q alpha-amylase.	4.68e+02
113	5	16.1	370 17	R91218	Human pituitary G-pro	4.68e+02	186	5	16.1	483 11	R58736	M15Q alpha-amylase.	4.68e+02
114	5	16.1	370 27	W31379	Human G protein-coupl	4.68e+02	187	5	16.1	483 11	R58004	M366Y alpha-amylase.	4.68e+02
115	5	16.1	372 30	W53184	Human GDF-1 (fx) prot	4.68e+02	188	5	16.1	483 11	R58001	M256Y alpha-amylase.	4.68e+02
116	5	16.1	372 10	R57980	Human GDF-1(fx) morph	4.68e+02	189	5	16.1	483 11	R58738	M15G alpha-amylase.	4.68e+02
117	5	16.1	372 9	R46750	Human morphogenic pro	4.68e+02	190	5	16.1	483 11	R58739	M15Y alpha-amylase.	4.68e+02
118	5	16.1	372 7	R34582	Human GDF-1 morphogen	4.68e+02	191	5	16.1	483 33	W57980	Alpha-amylase mutant	4.68e+02
119	5	16.1	372 9	R47260	Pre-pro-GDF-1.	4.68e+02	192	5	16.1	483 36	W80194	Bacillus licheniformi	4.68e+02
120	5	16.1	372 4	R23276	hGDF-1.	4.68e+02	193	5	16.1	483 2	R10579	Mutant alpha-amylase	4.68e+02
121	5	16.1	372 10	R50207	Human GDF-1(fx) morph	4.68e+02	194	5	16.1	483 33	W57977	Alpha-amylase mutant	4.68e+02
122	5	16.1	372 29	W40199	Human GDF-1 protein.	4.68e+02	195	5	16.1	483 33	W57979	Alpha-amylase mutant	4.68e+02
123	5	16.1	372 32	W48922	Human GDF-1.	4.68e+02	196	5	16.1	483 11	R58744	W138Y alpha-amylase.	4.68e+02
124	5	16.1	372 7	R33942	Morphogen GDF-1 full	4.68e+02	197	5	16.1	483 11	R58014	M197G alpha-amylase.	4.68e+02
125	5	16.1	372 26	W36862	Full length sequence	4.68e+02	198	5	16.1	483 11	R58733	W138C alpha-amylase.	4.68e+02
126	5	16.1	372 9	R47297	Human GDF-1(fx) morph	4.68e+02	199	5	16.1	483 28	W39742	B. licheniformis: alph	4.68e+02
127	5	16.1	372 10	R60961	Human islet cell anti	4.68e+02	200	5	16.1	483 2	R10577	Mutant alpha-amylase	4.68e+02
128	5	16.1	376 26	W35302	Consensus human hypot	4.68e+02	201	5	16.1	483 11	R57986	M197T alpha-amylase.	4.68e+02
129	5	16.1	380 24	W27510	Fragmented human NF-H	4.68e+02	202	5	16.1	483 11	R58015	M197Q alpha-amylase.	4.68e+02
130	5	16.1	386 22	W18664	Human ATP-sensitive p	4.68e+02	203	5	16.1	483 6	R34290	Temp-sensitive B.lich	4.68e+02
131	5	16.1	390 35	W74517	Human ATP-sensitive p	4.68e+02	204	5	16.1	483 36	W80214	Bacillus licheniformi	4.68e+02
132	5	16.1	390 35	W74516	Mouse ATP-sensitive p	4.68e+02	205	5	16.1	483 36	W80214	Mature alpha-amylase	4.68e+02
133	5	16.1	390 22	W17931	Mouse K-ATP channel s	4.68e+02	206	5	16.1	483 36	W80188	Bacillus licheniformi	4.68e+02
134	5	16.1	390 21	W13102	Human ATP sensitive p	4.68e+02	207	5	16.1	483 36	W80192	Bacillus licheniformi	4.68e+02
135	5	16.1	390 21	W13103	Mouse ATP sensitive p	4.68e+02	208	5	16.1	483 11	R58740	M15F alpha-amylase.	4.68e+02
136	5	16.1	390 22	W17932	Human K-ATP channel s	4.68e+02	209	5	16.1	483 11	R58746	W138L alpha-amylase.	4.68e+02
137	5	16.1	392 31	W53839	Pseudomonas Orfy secr	4.68e+02	210	5	16.1	483 18	R81477	Mutant mature alpha a	4.68e+02
138	5	16.1	396 33	W55070	Streptococcus pneumon	4.68e+02	211	5	16.1	483 18	R81475	Wild type mature alph	4.68e+02
139	5	16.1	410 2	R11546	T. hyo 39 KD family 2	4.68e+02	212	5	16.1	483 22	W08204	B.licheniformis mutan	4.68e+02
140	5	16.1	415 34	W69228	Human lysosomal siali	4.68e+02	213	5	16.1	483 22	W08203	B.licheniformis mutan	4.68e+02
141	5	16.1	419 28	W26744	S. carnosus nitrate r	4.68e+02	214	5	16.1	483 17	R88439	Alpha-amylase H133Y/T	4.68e+02
142	5	16.1	423 2	R11550	T. hyo gene 6:lac2 fu	4.68e+02	215	5	16.1	483 17	R88440	Alpha-amylase H133Y/T	4.68e+02
143	5	16.1	426 2	P70203	Sequence of human inh	4.68e+02	216	5	16.1	483 17	R88446	Alpha-amylase S187D m	4.68e+02
144	5	16.1	426 1	R05413	BUF-3 human different	4.68e+02	217	5	16.1	483 17	R88447	Alpha-amylase H133Y/S	4.68e+02
145	5	16.1	431 39	W87890	Protein encoded by th	4.68e+02	218	5	16.1	483 17	R88438	B. licheniformis matu	4.68e+02
146	5	16.1	431 1	R07464	Polypeptide with enzy	4.68e+02	219	5	16.1	483 17	R88445	Alpha-amylase V128E/H	4.68e+02
147	5	16.1	436 36	W80619	S. pneumoniae multisp	4.68e+02	220	5	16.1	483 11	R58742	M15Y alpha-amylase.	4.68e+02
148	5	16.1	450 2	R11551	T. hyo gene 6:lac2 fu	4.68e+02	221	5	16.1	483 22	W08202	B.licheniformis mutan	4.68e+02
149	5	16.1	452 15	R89280	CM-cellulase.	4.68e+02	222	5	16.1	483 2	R10578	Mutant alpha-amylase	4.68e+02
150	5	16.1	452 9	R49102	Translated sequence o	4.68e+02	223	5	16.1	483 4	R27250	M197F alpha-amylase.	4.68e+02
151	5	16.1	458 20	W09020	Neuronal nicotinic ac	4.68e+02	224	5	16.1	483 11	R58008	M197P alpha-amylase.	4.68e+02
152	5	16.1	461 23	W10210	Mature endoglycoceram	4.68e+02	225	5	16.1	483 11	R58011	B.licheniformis alpha	4.68e+02
153	5	16.1	471 8	R43646	Environmental adaptiv	4.68e+02	226	5	16.1	483 6	R31362	Bacillus licheniformi	4.68e+02
154	5	16.1	474 2	P70509	TrpC enzyme.	4.68e+02	227	5	16.1	483 22	W05229	Bacillus licheniformi	4.68e+02
155	5	16.1	482 22	W08195	B.licheniformis mutan	4.68e+02	228	5	16.1	483 22	W05232	Bacillus licheniformi	4.68e+02
156	5	16.1	482 22	W08196	B.licheniformis mutan	4.68e+02	229	5	16.1	483 17	R88441	Alpha-amylase N104D m	4.68e+02
157	5	16.1	482 22	W08201	B.licheniformis mutan	4.68e+02	230	5	16.1	483 11	R58734	M15Y alpha-amylase.	4.68e+02
158	5	16.1	482 22	W08192	B.licheniformis mutan	4.68e+02	231	5	16.1	483 11	R58737	M15E alpha-amylase.	4.68e+02
159	5	16.1	482 22	W08198	B.licheniformis mutan	4.68e+02	232	5	16.1	483 22	W08212	B.licheniformis mutan	4.68e+02
160	5	16.1	482 22	W08199	B.licheniformis mutan	4.68e+02	233	5	16.1	483 22	W08209	B.licheniformis mutan	4.68e+02
161	5	16.1	482 22	W08200	B.licheniformis mutan	4.68e+02	234	5	16.1	483 11	R58002	M304L alpha-amylase	4.68e+02
162	5	16.1	482 22	W08197	B.licheniformis mutan	4.68e+02	235	5	16.1	483 33	W57973	Alpha-amylase mutant	4.68e+02

236	5	16.1	483 33	W57975	Alpha-amylase mutant	4.68e+02	309	5	16.1	487 22	W05234	Bacillus licheniformi	4.68e+02
237	5	16.1	483 11	R58743	W138F alpha-amylase.	4.68e+02	310	5	16.1	489 20	W06824	Turkey herpes virus g	4.68e+02
238	5	16.1	483 4	R25434	Alpha-amylase mutant	4.68e+02	311	5	16.1	489 21	P90804	Glycoprotein A antige	4.68e+02
239	5	16.1	483 22	W08210	B.licheniformis mutan	4.68e+02	312	5	16.1	490 23	W10209	Full length endoglyco	4.68e+02
240	5	16.1	483 22	W08208	B.licheniformis mutan	4.68e+02	313	5	16.1	493 18	R91445	Human epithelial CD44	4.68e+02
241	5	16.1	483 11	R58723	M197W alpha-amylase.	4.68e+02	314	5	16.1	493 27	W20065	Human GADII protein.	4.68e+02
242	5	16.1	483 11	R58720	M197D alpha-amylase.	4.68e+02	315	5	16.1	493 24	W22306	Human GADII.	4.68e+02
243	5	16.1	483 11	R58730	M15K alpha-amylase.	4.68e+02	316	5	16.1	493 4	R20817	Epithelial CD44 Antig	4.68e+02
244	5	16.1	483 33	W57978	Alpha-amylase mutant	4.68e+02	317	5	16.1	499 37	R76296	Polyporus pinsitus (I	4.68e+02
245	5	16.1	483 17	R89448	Alpha-amylase H133Y/T	4.68e+02	318	5	16.1	499 37	W76299	Polyporus pinsitus (I	4.68e+02
246	5	16.1	483 18	R81476	Mutant mature alpha a	4.68e+02	319	5	16.1	499 37	W76297	Polyporus pinsitus (I	4.68e+02
247	5	16.1	483 17	R88442	Alpha-amylase N104D/H	4.68e+02	320	5	16.1	499 37	W76308	Polyporus pinsitus (I	4.68e+02
248	5	16.1	483 17	R88443	Alpha-amylase N104D/H	4.68e+02	321	5	16.1	499 34	W60875	Polyporus pinsitus (I	4.68e+02
249	5	16.1	483 17	R58745	W138A alpha-amylase.	4.68e+02	322	5	16.1	499 34	W60876	Polyporus pinsitus (I	4.68e+02
250	5	16.1	483 17	R88444	Alpha-amylase V128E/H	4.68e+02	323	5	16.1	499 37	W76295	Polyporus pinsitus (I	4.68e+02
251	5	16.1	483 11	R58007	M197I alpha-amylase.	4.68e+02	324	5	16.1	499 1	R07068	Phenol oxidase (PO) g	4.68e+02
252	5	16.1	483 11	R58006	M197A alpha-amylase.	4.68e+02	325	5	16.1	499 1	R07069	Phenol oxidase (PO) g	4.68e+02
253	5	16.1	483 33	W57971	Alpha-amylase mutant	4.68e+02	326	5	16.1	499 1	R05279	Amino acid sequence o	4.68e+02
254	5	16.1	483 33	W57972	Alpha-amylase mutant	4.68e+02	327	5	16.1	502 29	W45453	Arabidopsis thaliana	4.68e+02
255	5	16.1	483 11	R58721	M197E alpha-amylase.	4.68e+02	328	5	16.1	503 3	R14768	Metastasis-specific v	4.68e+02
256	5	16.1	483 33	W57982	Alpha-amylase mutant	4.68e+02	329	5	16.1	511 2	P70339	Alpha-amylase gene pr	4.68e+02
257	5	16.1	483 22	W05233	Bacillus licheniformi	4.68e+02	330	5	16.1	511 28	W39741	B. licheniformis alph	4.68e+02
258	5	16.1	483 22	W08194	B.licheniformis mutan	4.68e+02	331	5	16.1	511 29	W46865	Bacillus thuringiens	4.68e+02
259	5	16.1	483 11	R58735	M15L alpha-amylase.	4.68e+02	332	5	16.1	512 38	W73509	Alpha-amylase protein	4.68e+02
260	5	16.1	483 11	R58012	M197Y alpha-amylase.	4.68e+02	333	5	16.1	512 20	W10325	B. licheniformis acid	4.68e+02
261	5	16.1	483 33	W57981	Alpha-amylase mutant	4.68e+02	334	5	16.1	512 2	P70753	pheA arof amy operon	4.68e+02
262	5	16.1	483 11	R58731	M15P alpha-amylase.	4.68e+02	335	5	16.1	515 15	R90722	Laccase-LCC1.	4.68e+02
263	5	16.1	483 11	R58005	M438A alpha-amylase.	4.68e+02	336	5	16.1	520 15	R90721	Laccase-LCC2.	4.68e+02
264	5	16.1	483 18	R81478	Mutant mature alpha a	4.68e+02	337	5	16.1	525 10	R51272	Antigenic/functional	4.68e+02
265	5	16.1	483 18	R81480	Mutant mature alpha a	4.68e+02	338	5	16.1	531 18	R98007	PelB signal-Termamyl-	4.68e+02
266	5	16.1	483 11	R58003	M366A alpha-amylase.	4.68e+02	339	5	16.1	547 35	W1526	Helicobacter polyepet	4.68e+02
267	5	16.1	483 22	W05228	Bacillus licheniformi	4.68e+02	340	5	16.1	549 29	W55692	H. pylori ORF 06g205	4.68e+02
268	5	16.1	483 14	R72447	Bacillus licheniformi	4.68e+02	341	5	16.1	553 10	R54628	Manuronan C-5-epimer	4.68e+02
269	5	16.1	483 20	W12140	Alpha-amylase variant	4.68e+02	342	5	16.1	560 22	W20409	H. pylori secreted or	4.68e+02
270	5	16.1	483 11	R57919	M197K alpha-amylase.	4.68e+02	343	5	16.1	560 25	W22522	Alpha-amylase-cellulo	4.68e+02
271	5	16.1	483 11	R57985	Wild type alpha-amyla	4.68e+02	344	5	16.1	560 25	W23602	Alpha-amylase-cellulo	4.68e+02
272	5	16.1	483 33	W57974	Alpha-amylase mutant	4.68e+02	345	5	16.1	574 32	W57434	Termamyl-linker-CBD f	4.68e+02
273	5	16.1	483 21	W14498	Bacillus licheniformi	4.68e+02	346	5	16.1	577 8	R39703	Haemagglutinin-neura	4.68e+02
274	5	16.1	483 17	R88449	M15N alpha-amylase.	4.68e+02	347	5	16.1	577 20	W06827	Newcastle disease vir	4.68e+02
275	5	16.1	483 17	R88449	Alpha-amylase H133Y/N	4.68e+02	348	5	16.1	577 14	R80558	Newcastle disease vir	4.68e+02
276	5	16.1	483 20	W12139	Alpha-amylase variant	4.68e+02	349	5	16.1	581 11	R58598	Newcastle disease vir	4.68e+02
277	5	16.1	483 11	R57992	M15L alpha-amylase.	4.68e+02	350	5	16.1	581 11	R58599	Newcastle disease vir	4.68e+02
278	5	16.1	483 11	R58010	M197S alpha-amylase.	4.68e+02	351	5	16.1	581 21	W10690	Newcastle disease vir	4.68e+02
279	5	16.1	483 17	R88450	Alpha-amylase H133Y/T	4.68e+02	352	5	16.1	585 38	W73315	Parathyroid hormone r	4.68e+02
280	5	16.1	483 11	R58009	M197V alpha-amylase.	4.68e+02	353	5	16.1	585 5	R27705	Opossum kidney PTH/PT	4.68e+02
281	5	16.1	483 22	W08193	B.licheniformis mutan	4.68e+02	354	5	16.1	585 16	R92276	Opossum kidney PTH/PT	4.68e+02
282	5	16.1	483 22	W05230	Bacillus licheniformi	4.68e+02	355	5	16.1	589 33	W54096	Homo sapiens TGL163 s	4.68e+02
283	5	16.1	483 22	W08207	B.licheniformis mutan	4.68e+02	356	5	16.1	593 5	R27220	Brain GAD #2.	4.68e+02
284	5	16.1	483 22	W08211	B.licheniformis mutan	4.68e+02	357	5	16.1	594 36	W74717	Amino acid sequence o	4.68e+02
285	5	16.1	483 20	W12137	Alpha-amylase variant	4.68e+02	358	5	16.1	594 36	W74716	Amino acid sequence o	4.68e+02
286	5	16.1	483 22	W08206	Alpha-amylase variant	4.68e+02	359	5	16.1	594 5	R27221	Full length brain GAD	4.68e+02
287	5	16.1	483 11	R58724	M197R alpha-amylase.	4.68e+02	360	5	16.1	594 5	R27222	Full length islet GAD	4.68e+02
288	5	16.1	483 11	R58726	M15R alpha-amylase.	4.68e+02	361	5	16.1	599 23	W17788	Phage abortive infect	4.68e+02
289	5	16.1	483 20	W12138	Alpha-amylase variant	4.68e+02	362	5	16.1	607 32	W60268	Klebsiella pneumoniae	4.68e+02
290	5	16.1	483 20	W12137	Alpha-amylase variant	4.68e+02	363	5	16.1	616 1	R06329	Newcastle disease vir	4.68e+02
291	5	16.1	483 22	W08205	B.licheniformis mutan	4.68e+02	364	5	16.1	616 2	P96147	Sequence of Newcastle	4.68e+02
292	5	16.1	483 11	R58728	M15D alpha-amylase.	4.68e+02	365	5	16.1	621 29	W55645	H. pylori ORF 06ap106	4.68e+02
293	5	16.1	485 27	W31500	Bacillus sp. alpha am	4.68e+02	366	5	16.1	630 25	W22523	Alpha-amylase-CenA ce	4.68e+02
294	5	16.1	485 30	W48261	Bacillus sp. alpha am	4.68e+02	367	5	16.1	630 25	W23603	Alpha-amylase-CenA ce	4.68e+02
295	5	16.1	485 16	R81836	Bacillus sp. alkaline	4.68e+02	368	5	16.1	637 30	W57539	H. pylori ORF 06ep106	4.68e+02
296	5	16.1	485 20	W12956	Alpha-amylase.	4.68e+02	369	5	16.1	638 5	R03924	E. coli HSP (dnak)	4.68e+02
297	5	16.1	487 36	W08215	A4 form alpha-amylase	4.68e+02	370	5	16.1	641 26	W35298	Macaque islet cell an	4.68e+02
298	5	16.1	487 36	W80193	A4 form alpha-amylase	4.68e+02	371	5	16.1	653 32	W57437	NAM1 0 fusion constru	4.68e+02
299	5	16.1	487 22	W05239	Bacillus licheniformi	4.68e+02	372	5	16.1	657 6	R28964	Notch hNSK full lengt	4.68e+02
300	5	16.1	487 11	R57987	A4 form alpha-amylase	4.68e+02	373	5	16.1	658 17	R58586	WD-40 domain-contg. C	4.68e+02
301	5	16.1	487 36	W80213	A4 form alpha-amylase	4.68e+02	374	5	16.1	714 27	W35316	M. cattarrhalis RI tra	4.68e+02
302	5	16.1	487 36	W80190	Mutant of alpha-amyla	4.68e+02	375	5	16.1	717 19	W06109	Scytalidium catalase.	4.68e+02
303	5	16.1	487 36	W80189	A4 form of the alpha-	4.68e+02	376	5	16.1	721 29	W55409	H. pylori ORF 14ep119	4.68e+02
304	5	16.1	487 36	W80191	A4 form alpha-amylase	4.68e+02	377	5	16.1	730 31	W46315	Human elastin contain	4.68e+02
305	5	16.1	487 22	W05235	Bacillus licheniformi	4.68e+02	378	5	16.1	733 11	R56653	Synthetic human tropo	4.68e+02
306	5	16.1	487 22	W05237	Bacillus licheniformi	4.68e+02	379	5	16.1	739 8	R43032	Human complement fact	4.68e+02
307	5	16.1	487 22	W05238	Bacillus licheniformi	4.68e+02	380	5	16.1	746 37	W80322	Bacillus thuringiens	4.68e+02
308	5	16.1	487 22	W05236	Bacillus licheniformi	4.68e+02	381	5	16.1	746 16	R92693	Isoamylase mature pro	4.68e+02

382	5	16.1	746 30	W53886	Mature Flavobacterium	4.68e+02	455	5	16.1	1288 29	W55547	H. pylori ORF 14ee119	4.68e+02
383	5	16.1	750 25	W33600	Alpha-amylase-cellulo	4.68e+02	456	5	16.1	1288 29	W55685	H. pylori ORF 07ee114	4.68e+02
384	5	16.1	752 25	W22520	Alpha-amylase-cellulo	4.68e+02	457	5	16.1	1296 8	R41198	CT	4.68e+02
385	5	16.1	752 23	W9036	Mismatch repair prote	4.68e+02	458	5	16.1	1297 30	W40200	Infected cell protein	4.68e+02
386	5	16.1	756 14	R76071	Human mismatch repair	4.68e+02	459	5	16.1	1298 37	W80810	The amino acid sequen	4.68e+02
387	5	16.1	756 15	R75785	Human wild type MLH1,	4.68e+02	460	5	16.1	1403 10	R54629	Mannuronan C-5-epimer	4.68e+02
388	5	16.1	756 15	R79008	Human DNA repair prot,	4.68e+02	461	5	16.1	1413 22	W20725	H. pylori secreted or	4.68e+02
389	5	16.1	759 29	W46864	Bacillus thuringiens	4.68e+02	462	5	16.1	1580 26	W22619	Platenolide synthase	4.68e+02
390	5	16.1	775 15	R79950	Enzyme Q36	4.68e+02	463	5	16.1	1580 26	W23719	Platenolide synthase	4.68e+02
391	5	16.1	777 30	W33885	Full length Flavobact	4.68e+02	464	5	16.1	1612 32	W65088	R. prowazekii S-layer	4.68e+02
392	5	16.1	777 16	R92692	Isoamylase precursor	4.68e+02	465	5	16.1	1723 19	W00645	Mouse DEC-205	4.68e+02
393	5	16.1	782 32	W47336	Teramylase precursor	4.68e+02	466	5	16.1	1829 27	W29322	DNA polymerase with 3	4.68e+02
394	5	16.1	789 29	W46726	Native vegetative ins	4.68e+02	467	5	16.1	2353 19	R99393	Haemophilus adhesio	4.68e+02
395	5	16.1	789 17	R91244	B. thuringiensis VIP3	4.68e+02	468	5	16.1	2756 30	W37050	S. putrefaciens EPO b	4.68e+02
396	5	16.1	789 37	W80320	Bacillus thuringiens	4.68e+02	469	5	16.1	2756 20	R99462	Biosynthetic enzyme o	4.68e+02
397	5	16.1	789 29	W46869	Bacillus thuringiens	4.68e+02	470	5	16.1	2763 20	W10344	Maize dwarf mosaic vi	4.68e+02
398	5	16.1	789 37	W80321	Bacillus thuringiens	4.68e+02	471	5	16.1	2864 16	R82072	Hepatitis GB virus (H	4.68e+02
399	5	16.1	789 32	W60217	Bacillus thuringiens	4.68e+02	472	5	16.1	3054 8	R40841	Translaton of TEV la	4.68e+02
400	5	16.1	789 17	R91243	B. thuringiensis VIP3	4.68e+02	473	5	16.1	3110 15	R71730	Merotin major subunit	4.68e+02
401	5	16.1	789 29	W46868	Bacillus thuringiens	4.68e+02	474	5	16.1	3567 8	R44431	erva region polypepti	4.68e+02
402	5	16.1	789 29	W46867	Bacillus thuringiens	4.68e+02	475	5	16.1	3722 2	R10145	Cephalosporin antibio	4.68e+02
403	5	16.1	789 29	W46861	Bacillus thuringiens	4.68e+02	476	5	16.1	4472 26	W22601	Tyactone synthase OR	4.68e+02
404	5	16.1	789 29	W46862	Bacillus thuringiens	4.68e+02	477	4	12.9	6 24	W26334	Silk fibroin repeat u	3.41e+03
405	5	16.1	789 29	W46825	Vegetative insecticid	4.68e+02	478	4	12.9	6 19	W07532	Natural fibroin deriv	3.41e+03
406	5	16.1	789 29	W46866	Bacillus thuringiens	4.68e+02	479	4	12.9	9 21	W13092	Plasmodium vivax circ	3.41e+03
407	5	16.1	789 29	W46863	Bacillus thuringiens	4.68e+02	480	4	12.9	16 39	W89310	Human rchd534 peptid	3.41e+03
408	5	16.1	789 28	W9515	B. cereus VIP2A(b) pr	4.68e+02	481	4	12.9	17 39	W89159	Anti-p53 monoclonal a	3.41e+03
409	5	16.1	789 28	W9514	B. cereus VIP2A(a) pr	4.68e+02	482	4	12.9	17 5	R25056	Synthetic random copo	3.41e+03
410	5	16.1	789 32	W60216	Bacillus thuringiens	4.68e+02	483	4	12.9	19 3	P60474	Epstein-Barr virus nu	3.41e+03
411	5	16.1	789 29	W46860	Bacillus thuringiens	4.68e+02	484	4	12.9	19 24	W30153	Tandem SH2 ligand Fee	3.41e+03
412	5	16.1	789 29	W46870	Bacillus thuringiens	4.68e+02	485	4	12.9	20 24	W27024	Monomer subunit of er	3.41e+03
413	5	16.1	790 29	W46872	Bacillus thuringiens	4.68e+02	486	4	12.9	21 5	R25054	Synthetic random copo	3.41e+03
414	5	16.1	790 32	W60215	Bacillus thuringiens	4.68e+02	487	4	12.9	22 18	W00761	PACAP2-23-NH2	3.41e+03
415	5	16.1	790 29	W46871	Bacillus thuringiens	4.68e+02	488	4	12.9	23 17	R90188	PACAP23 for promotng	3.41e+03
416	5	16.1	799 28	W23938	Wheat granule-bound s	4.68e+02	489	4	12.9	23 21	W14084	Streptococcus bovis m	3.41e+03
417	5	16.1	800 26	W34479	Zfx protein	4.68e+02	490	4	12.9	23 21	W15755	Cyclic adenylate mono	3.41e+03
418	5	16.1	809 17	R46732	Maize optimised VIP3A	4.68e+02	491	4	12.9	25 39	W95575	Analog omega-conopept	3.41e+03
419	5	16.1	809 17	R91261	VIP3A(a) synthetic:na	4.68e+02	492	4	12.9	25 39	W95574	Analog omega-conopept	3.41e+03
420	5	16.1	809 28	W9521	B. cereus VIP2A(a) sy	4.68e+02	493	4	12.9	25 39	W95577	Analog omega-conopept	3.41e+03
421	5	16.1	818 26	W35297	Human islet cell anti	4.68e+02	494	4	12.9	25 39	W95578	Analog omega-conopept	3.41e+03
422	5	16.1	872 10	R34631	Fragment of mannurona	4.68e+02	495	4	12.9	25 39	W95582	Analog omega-conopept	3.41e+03
423	5	16.1	876 23	W18091	Type I diabetes-asso	4.68e+02	496	4	12.9	25 10	R51771	Der p II derived pept	3.41e+03
424	5	16.1	906 35	W71290	Potato starch branchi	4.68e+02	497	4	12.9	26 39	W95576	Analog omega-conopept	3.41e+03
425	5	16.1	906 35	W69300	Potato class B starch	4.68e+02	498	4	12.9	26 18	W04798	PACAP(1-26)-NH2 for t	3.41e+03
426	5	16.1	912 32	W42103	Human p115 Rho-guanin	4.68e+02	499	4	12.9	27 38	W89019	Sequence ID #719 from	3.41e+03
427	5	16.1	915 31	W23306	Human pro-protein con	4.68e+02	500	4	12.9	27 39	W95580	Analog omega-conopept	3.41e+03
428	5	16.1	937 2	R10695	Cephalosporin antibio	4.68e+02	501	4	12.9	27 39	W95579	Analog omega-conopept	3.41e+03
429	5	16.1	969 24	W25170	Human insulinoma-asso	4.68e+02	502	4	12.9	27 6	R32722	Argl17-PACAP27-NH2	3.41e+03
430	5	16.1	986 24	W25171	Human insulinoma-asso	4.68e+02	503	4	12.9	27 18	W03822	[Glu3]-PACAP(1-27)-NH	3.41e+03
431	5	16.1	997 10	R34630	Mannuronan C-5-epimer	4.68e+02	504	4	12.9	28 32	W60195	Mitochondria Pol I-ty	3.41e+03
432	5	16.1	999 6	R30742	Human pemphigus vulga	4.68e+02	505	4	12.9	29 39	W04664	[Glu3 8]-PACAP(1-29)-	3.41e+03
433	5	16.1	1012 26	W35296	Macaque islet cell an	4.68e+02	506	4	12.9	30 39	W89258	Human PTP04 peptide 4	3.41e+03
434	5	16.1	1015 23	R18092	Type I diabetes-asso	4.68e+02	507	4	12.9	30 18	W03770	PACAP(1-30)-NH2 for t	3.41e+03
435	5	16.1	1015 27	W35345	Human protein tyrosin	4.68e+02	508	4	12.9	30 17	R90195	PACAP30 for promotng	3.41e+03
436	5	16.1	1026 9	R48993	rsaA S-layer protein.	4.68e+02	509	4	12.9	30 3	R14368	N-terminal of spider	3.41e+03
437	5	16.1	1026 16	R34014	Caulobacter S-layer p	4.68e+02	510	4	12.9	30 3	R12631	[Ser-Thr]2-Lys7]-Avi	3.41e+03
438	5	16.1	1026 27	W37490	Caulobacter crescentu	4.68e+02	511	4	12.9	31 17	R90196	PACAP31 for promotng	3.41e+03
439	5	16.1	1049 25	W22159	ApxIIb protein.	4.68e+02	512	4	12.9	31 10	R53482	CD44 peptide CD44-9	3.41e+03
440	5	16.1	1058 9	R42453	Enzyme involved in ei	4.68e+02	513	4	12.9	32 38	W75416	Bovine alpha-endothel	3.41e+03
441	5	16.1	1096 20	W01596	Inositol-1,4,5-tripho	4.68e+02	514	4	12.9	33 39	W90033	B. sphaericus phenyla	3.41e+03
442	5	16.1	1098 19	R93932	Haemophilus adhesio	4.68e+02	515	4	12.9	33 17	R90198	PACAP33 for promotng	3.41e+03
443	5	16.1	1130 3	R13436	Merotin M polypeptide	4.68e+02	516	4	12.9	33 18	W04718	[Ser24,25]-PACAP(1-35	3.41e+03
444	5	16.1	1130 15	R71729	Merotin major subunit	4.68e+02	517	4	12.9	36 14	R75172	Guayule rubber partic	3.41e+03
445	5	16.1	1157 19	W06418	Antiscarab pest toxin	4.68e+02	518	4	12.9	37 17	R93976	PACAP-38 for ameliora	3.41e+03
446	5	16.1	1157 6	R28900	Toxin 50C	4.68e+02	519	4	12.9	38 18	W04721	[Ser24,25]-PACAP(1-38	3.41e+03
447	5	16.1	1157 5	R27343	B. thuringiensis PS50C	4.68e+02	520	4	12.9	38 18	W00737	NaiphaAcetyl/Naipha-S	3.41e+03
448	5	16.1	1157 14	R44208	Bacillus thuringiens	4.68e+02	521	4	12.9	38 2	R10006	Pituitary adenylate c	3.41e+03
449	5	16.1	1157 7	R33768	Bt isolate Ps50C	4.68e+02	522	4	12.9	39 38	W82464	Onchocerca sp Cu/Zn S	3.41e+03
450	5	16.1	1157 5	R25997	Delta-endotoxin.	4.68e+02	523	4	12.9	46 31	W27944	Staphylococcus aureus	3.41e+03
451	5	16.1	1169 19	W06417	Antiscarab pest toxin	4.68e+02	524	4	12.9	48 38	W89076	Polypeptide fragment	3.41e+03
452	5	16.1	1239 31	W36131	Leukothia melanogaste	4.68e+02	525	4	12.9	48 32	W28295	Staphylococcus aureus	3.41e+03
453	5	16.1	1244 10	R34781	Leukothia AppiiaA	4.68e+02	526	4	12.9	50 39	W89337	Pasturella haemolytic	3.41e+03
454	5	16.1	1287 15	R79944	Helicobacter pylori v	4.68e+02	527	4	12.9	52 38	W75417	Bovine beta-endotheli	3.41e+03



528	4	12.9	53 31	W53509	Synthetic peptide 1.	3.41e+03	601	186 39	W95507	Deduced amino acid se	3.41e+03
529	4	12.9	56 2	P71025	Sequence of portion o	3.41e+03	602	186 38	W85120	Desaturase enzyme par	3.41e+03
530	4	12.9	58 39	W83327	MP47 peptide sequence	3.41e+03	603	191 38	W85579	Human N-methyl-D-aspa	3.41e+03
531	4	12.9	59 33	W61320	Mutant BCL-XL/BCL-2 a	3.41e+03	604	196 3	P60795	Aquarin protein.	3.41e+03
532	4	12.9	61 2	P1076	Sequence encoded by h	3.41e+03	605	199 10	W53292	IgG-binding Streptoco	3.41e+03
533	4	12.9	69 30	W27791	Staphylococcus aureus	3.41e+03	606	199 1	P94783	Protein G variant.	3.41e+03
534	4	12.9	71 26	W23068	Canine IgE heavy chai	3.41e+03	607	200 31	W28317	Amino acid sequence o	3.41e+03
535	4	12.9	72 21	W11799	Onchocerca volvulus L	3.41e+03	608	204 39	W95361	Guinea pig interleuki	3.41e+03
536	4	12.9	89 39	W90047	B. steeraotherophilus	3.41e+03	609	204 30	W58832	Murine BAG protein.	3.41e+03
537	4	12.9	89 22	W20444	H. pylori cytoplasmic	3.41e+03	610	204 33	W61318	Mutant BCL-XL/BCL-2 a	3.41e+03
538	4	12.9	91 38	W85914	Amino acid sequence o	3.41e+03	611	206 38	W85097	Thyroid hormone recep	3.41e+03
539	4	12.9	91 38	W85906	PrP 81 heavy chain va	3.41e+03	612	206 38	W88747	Secreted protein enco	3.41e+03
540	4	12.9	92 38	W85921	Amino acid sequence o	3.41e+03	613	208 39	W89752	Staphylococcus aureus	3.41e+03
541	4	12.9	92 38	W85915	Amino acid sequence o	3.41e+03	614	210 38	W89867	Antigen 2 from cluste	3.41e+03
542	4	12.9	92 38	W88725	Secreted protein enco	3.41e+03	615	210 38	W89990	Expressed antigen for	3.41e+03
543	4	12.9	95 38	W85922	Amino acid sequence o	3.41e+03	616	216 39	W67858	Human secreted protei	3.41e+03
544	4	12.9	101 38	W85918	Amino acid sequence o	3.41e+03	617	218 5	R28645	UL26 protease deletio	3.41e+03
545	4	12.9	103 38	W85920	Amino acid sequence o	3.41e+03	618	219 5	R26540	Sequence of 40 kD ant	3.41e+03
546	4	12.9	103 14	R74928	Porcine elstatin cath	3.41e+03	619	225 39	W81346	Tomato expansin LeEx1	3.41e+03
547	4	12.9	105 38	W85911	Amino acid sequence o	3.41e+03	620	226 25	W33273	S. fradiae tylosin bi	3.41e+03
548	4	12.9	108 38	W89941	Antigen 2 from cluste	3.41e+03	621	228 38	W83146	B29 receptor protein.	3.41e+03
549	4	12.9	108 38	W85919	Amino acid sequence o	3.41e+03	622	229 38	W89855	Antigen 1 from cluste	3.41e+03
550	4	12.9	109 1	R07009	Protein G variant.	3.41e+03	623	232 38	W87489	S. cerevisiae TIH3 po	3.41e+03
551	4	12.9	111 39	W89172	Anti-p53 monoclonal a	3.41e+03	624	234 39	W89972	Antigen 1 from cluste	3.41e+03
552	4	12.9	113 38	W88602	Secreted protein enco	3.41e+03	625	234 38	W89826	Protein encoded by cl	3.41e+03
553	4	12.9	113 24	W25749	Human tctex-1.	3.41e+03	626	235 39	W89301	Human rchd534 protein	3.41e+03
554	4	12.9	114 38	W85912	Amino acid sequence o	3.41e+03	627	236 39	W89788	Staphylococcus aureus	3.41e+03
555	4	12.9	114 38	W85913	Amino acid sequence o	3.41e+03	628	236 39	W89985	Expressed antigen for	3.41e+03
556	4	12.9	114 38	W85904	PrP 28 heavy chain va	3.41e+03	629	237 39	W89447	A g1bB polypeptide se	3.41e+03
557	4	12.9	114 38	W85905	PrP 28 heavy chain va	3.41e+03	630	238 30	W52331	Engineered green fluo	3.41e+03
558	4	12.9	114 1	P94770	ayw Pre-S1 sequence o	3.41e+03	631	238 30	W52332	Engineered green fluo	3.41e+03
559	4	12.9	116 39	W81342	Rat cocaine and amphe	3.41e+03	632	238 32	W40479	A. victoria green flu	3.41e+03
560	4	12.9	117 32	W54451	Mouse novel secreted	3.41e+03	633	238 30	W52343	Engineered green fluo	3.41e+03
561	4	12.9	117 30	W37168	Cooperia onchophora ES	3.41e+03	634	238 30	W52335	Engineered green fluo	3.41e+03
562	4	12.9	119 1	P90540	Immunoglobulin H chai	3.41e+03	635	238 30	W23831	S. glaucescens acBA p	3.41e+03
563	4	12.9	123 39	W87627	Dirofilaria immitis a	3.41e+03	636	238 30	W52342	Engineered green fluo	3.41e+03
564	4	12.9	125 1	P94769	adw Pre-S1 sequence o	3.41e+03	637	238 30	W05306	Green fluorescent pro	3.41e+03
565	4	12.9	125 1	P80289	Human Vpre-B protein.	3.41e+03	638	238 32	W42492	A. victoria green flu	3.41e+03
566	4	12.9	126 5	R26075	Mutant E2 binding dom	3.41e+03	639	238 32	W42488	A. victoria green flu	3.41e+03
567	4	12.9	127 28	W27652	Secreted protein AS34	3.41e+03	640	238 30	W52326	Engineered green fluo	3.41e+03
568	4	12.9	129 39	W81330	Rat cocaine and amphe	3.41e+03	641	240 13	R68613	Single chain antibody	3.41e+03
569	4	12.9	129 22	W06813	Rat cocaine and amphe	3.41e+03	642	242 18	W00166	Meiosis specific prot	3.41e+03
570	4	12.9	130 15	R85118	Cell-cycle regulatory	3.41e+03	643	246 39	W89781	Staphylococcus aureus	3.41e+03
571	4	12.9	130 38	W70823	Mouse multiple tumour	3.41e+03	644	248 29	W37091	Lycepersicon esculent	3.41e+03
572	4	12.9	131 10	R53300	IgG-binding Streptoco	3.41e+03	645	250 39	W67745	Deleya halophila male	3.41e+03
573	4	12.9	138 15	R85117	Cell-cycle regulatory	3.41e+03	646	250 25	W35088	E. coli serine acetyl	3.41e+03
574	4	12.9	138 38	W80526	A human multiple tumo	3.41e+03	647	253 21	W11886	Hantavirus nuclear pr	3.41e+03
575	4	12.9	138 1	P90542	Amino acid sequence o	3.41e+03	648	253 38	W82637	Ehrlichia sp. HGE-7 p	3.41e+03
576	4	12.9	143 38	W89078	Polypeptide fragment	3.41e+03	649	256 17	R94719	PRRSV VR 2385 ORF-2 p	3.41e+03
577	4	12.9	145 1	P94865	Der p II antigen from	3.41e+03	650	256 25	W35086	E. coli serine acetyl	3.41e+03
578	4	12.9	148 1	P94782	Protein G variant.	3.41e+03	651	257 2	P70404	ORF 1 gene product of	3.41e+03
579	4	12.9	154 39	W90081	C. cellulovorans CBD-	3.41e+03	652	257 38	W88578	Secreted protein enco	3.41e+03
580	4	12.9	155 39	W89807	Staphylococcus aureus	3.41e+03	653	257 25	W35085	E. coli serine acetyl	3.41e+03
581	4	12.9	156 39	W90080	C. cellulovorans CBD-	3.41e+03	654	258 3	P50191	Sequence encoded by t	3.41e+03
582	4	12.9	156 22	W20959	H. pylori cytoplasmic	3.41e+03	655	260 39	W95531	Carbohydrate-binding	3.41e+03
583	4	12.9	159 38	W82486	Ehrlichia sp. E74.1 p	3.41e+03	656	264 38	W82268	Human NTN-2 protein.	3.41e+03
584	4	12.9	159 38	W88799	Polypeptide fragment	3.41e+03	657	266 33	W62462	Human T cell surface	3.41e+03
585	4	12.9	161 27	W31219	HIV-TAT protein trans	3.41e+03	658	269 21	W21139	Lipolytic enzyme, D13	3.41e+03
586	4	12.9	162 39	W90077	C. cellulovorans CbpA	3.41e+03	659	269 21	W21118	Lipolytic enzyme, D57	3.41e+03
587	4	12.9	163 21	W11280	Human breast specific	3.41e+03	660	269 21	W21117	Lipolytic enzyme, D57	3.41e+03
588	4	12.9	164 10	R53299	IgG-binding Streptoco	3.41e+03	661	269 21	W53295	IgG-binding Streptoco	3.41e+03
589	4	12.9	166 38	W73468	Human secreted protei	3.41e+03	662	269 14	R81098	Lipolytic enzyme, D57	3.41e+03
590	4	12.9	170 33	W63805	H. pylori strain J166	3.41e+03	663	269 14	R81993	Variant lipase of Hum	3.41e+03
591	4	12.9	170 39	W67874	Human secreted protei	3.41e+03	664	269 14	R82000	Variant lipase of Hum	3.41e+03
592	4	12.9	173 17	R94705	PRRSV Lelystad ORF-6	3.41e+03	665	269 1	P91444	Region of tryptophan	3.41e+03
593	4	12.9	174 3	W15618	HBSAg pre-S region su	3.41e+03	666	274 18	W00080	Alkaline protease APL	3.41e+03
594	4	12.9	176 39	W84069	Amino acid sequence o	3.41e+03	667	274 21	W15448	Subtilisin Carlsberg	3.41e+03
595	4	12.9	180 3	R41408	Zea mays 15 kD seed s	3.41e+03	668	274 21	W15453	Subtilisin Carlsberg	3.41e+03
596	4	12.9	181 39	W89199	Human SP-18 surfactan	3.41e+03	669	275 21	W21177	Lipolytic enzyme, El-S	3.41e+03
597	4	12.9	181 38	W89810	Protein encoded by cl	3.41e+03	670	277 39	W84054	Human v3 loop HIV rec	3.41e+03
598	4	12.9	183 9	R44491	Streptavidin gene.	3.41e+03	671	277 21	W21198	Lipolytic enzyme, El-S	3.41e+03
599	4	12.9	184 39	W67906	Human secreted protei	3.41e+03	672	277 21	W21160	Lipolytic enzyme, SPP	3.41e+03
600	4	12.9	185 31	W28009	Staphylococcus aureus	3.41e+03	673				

674	4	12.9	286 39	W88239	Streptococcus uberis	3.41e+03	747	4	12.9	400 1	P93189	E.coli acetate kinase	3.41e+03
675	4	12.9	286 39	W88240	Streptococcus uberis	3.41e+03	748	4	12.9	402 16	R83407	Humicola insolens end	3.41e+03
676	4	12.9	286 1	P82933	SAP(Phel)	3.41e+03	749	4	12.9	402 16	R83408	Humicola insolens end	3.41e+03
677	4	12.9	289 38	W70806	A human inorganic pyr	3.41e+03	750	4	12.9	402 2	R06495	Beta 3 adrenergic rec	3.41e+03
678	4	12.9	289 39	W67881	Human secreted protei	3.41e+03	751	4	12.9	408 39	W89442	Mycoplasma hominis ar	3.41e+03
679	4	12.9	289 23	W19647	Humicola lanuginosa m	3.41e+03	752	4	12.9	410 21	W12093	Bovine papillomavirus	3.41e+03
680	4	12.9	291 12	R65399	Variant lipase E56W,	3.41e+03	753	4	12.9	412 3	R15647	Sau3AI methylase	3.41e+03
681	4	12.9	291 12	R65425	Variant lipase L259F,	3.41e+03	754	4	12.9	413 22	W20759	H. pylori secreted or	3.41e+03
682	4	12.9	291 12	R65425	Variant lipase L259M,	3.41e+03	755	4	12.9	413 39	W89678	Clostridium chauvoei	3.41e+03
683	4	12.9	291 25	R32623	Lipase for use in a n	3.41e+03	756	4	12.9	413 27	W31660	Bovine papillomavirus	3.41e+03
684	4	12.9	291 4	R22632	Generic sequence of H	3.41e+03	757	4	12.9	417 31	W54358	Phosphoglycerate kina	3.41e+03
685	4	12.9	291 4	R22632	Humicola lanuginosa l	3.41e+03	758	4	12.9	417 39	W89243	Rhodococcus sp. strai	3.41e+03
686	4	12.9	291 12	R65429	Variant lipase L206K,	3.41e+03	759	4	12.9	418 39	W73575	GST-Msp I methylase f	3.41e+03
687	4	12.9	291 12	R65434	Variant lipase L206Q,	3.41e+03	760	4	12.9	418 39	W73574	GST-Msp I methylase f	3.41e+03
688	4	12.9	293 20	W06244	Apoptotic cysteine pr	3.41e+03	761	4	12.9	420 38	W68541	Amino acid sequence o	3.41e+03
689	4	12.9	297 21	W00282	Mutant farnesylidiphos	3.41e+03	762	4	12.9	420 22	W20266	H. pylori transmembra	3.41e+03
690	4	12.9	297 32	W62532	Farnesyl diphosphate s	3.41e+03	763	4	12.9	421 17	R97600	Adenovirus 11 subtype	3.41e+03
691	4	12.9	297 21	W00283	Mutant farnesylidiphos	3.41e+03	764	4	12.9	421 17	R97599	Adenovirus 8 subtype	3.41e+03
692	4	12.9	299 39	W80985	B cell receptor assoc	3.41e+03	765	4	12.9	421 33	W69547	Brevibacterium lactof	3.41e+03
693	4	12.9	299 39	W80984	Amino acid sequence o	3.41e+03	766	4	12.9	423 39	W73767	M. tuberculosis antig	3.41e+03
694	4	12.9	301 39	W87740	Cornebacterium dihyd	3.41e+03	767	4	12.9	423 39	W73657	M. tuberculosis antig	3.41e+03
695	4	12.9	301 33	W69549	Brevibacterium lactof	3.41e+03	768	4	12.9	423 17	R85876	WD-40 domain-contg. y	3.41e+03
696	4	12.9	302 11	R55793	Herpesvirus saimiri s	3.41e+03	769	4	12.9	424 22	W20977	H. pylori transmembra	3.41e+03
697	4	12.9	303 33	W62772	Human immunoglobulin	3.41e+03	770	4	12.9	432 38	W82246	Bacteriophage fd-tet	3.41e+03
698	4	12.9	306 38	W85132	A desaturase enzyme e	3.41e+03	771	4	12.9	434 3	R12533	Inosine guanosine kin	3.41e+03
699	4	12.9	306 38	W84153	Human desaturase enzy	3.41e+03	772	4	12.9	435 38	R12533	Ehrlichia sp. E8 prot	3.41e+03
700	4	12.9	306 39	W95511	Amino acid sequence o	3.41e+03	773	4	12.9	435 39	W89278	Granulocytic Ehrlich	3.41e+03
701	4	12.9	310 39	W95197	Human syndecan-1.	3.41e+03	774	4	12.9	436 24	W23218	Protein product of b.	3.41e+03
702	4	12.9	312 21	W09784	L-proline-3-hydroxyla	3.41e+03	775	4	12.9	437 31	W48389	Homo sapiens E2F1 tra	3.41e+03
703	4	12.9	317 38	W73507	Human ATG-1709 protei	3.41e+03	776	4	12.9	441 21	W09808	Platelet-activating f	3.41e+03
704	4	12.9	321 39	W67847	Human secreted protei	3.41e+03	777	4	12.9	441 38	W73359	Human PAF-AH protein	3.41e+03
705	4	12.9	324 6	R30971	TGF-beta-1 binding pr	3.41e+03	778	4	12.9	441 14	R73049	Human acetyl hydrolas	3.41e+03
706	4	12.9	325 38	W82615	Ehrlichia sp. HGE-1 p	3.41e+03	779	4	12.9	442 38	W85035	Green fluorescent pro	3.41e+03
707	4	12.9	325 4	P93704	Sequence of the 65kd	3.41e+03	780	4	12.9	443 26	W35744	Treponema pallidum 47	3.41e+03
708	4	12.9	326 38	W89279	Granulocytic Ehrlich	3.41e+03	781	4	12.9	444 38	W73362	Bovine PAF-AH protein	3.41e+03
709	4	12.9	326 38	W81983	Ehrlichia sp. E46.1 p	3.41e+03	782	4	12.9	446 39	W95506	Mortierella alpina de	3.41e+03
710	4	12.9	328 39	W67852	Human secreted protei	3.41e+03	783	4	12.9	446 38	W85119	A delta-5 desaturase	3.41e+03
711	4	12.9	330 38	W86020	Human homeobox gene H	3.41e+03	784	4	12.9	448 2	R07013	Protein G variant wit	3.41e+03
712	4	12.9	333 38	W86019	Murine homeobox Hox11	3.41e+03	785	4	12.9	459 39	W89182	S. pneumoniae Glnu po	3.41e+03
713	4	12.9	334 23	W19631	Human mitogen-activat	3.41e+03	786	4	12.9	461 39	W89558	Triticum sp. cysteine	3.41e+03
714	4	12.9	336 38	W86338	Kidney injury associa	3.41e+03	787	4	12.9	465 9	R47972	Lactobacillus brevis	3.41e+03
715	4	12.9	338 38	W82407	E. chrysanthemi HrpN	3.41e+03	788	4	12.9	467 8	R39470	MRX-alpha	3.41e+03
716	4	12.9	338 38	W87638	A hypersensitive resp	3.41e+03	789	4	12.9	469 39	W89560	Triticum sp. cysteine	3.41e+03
717	4	12.9	338 20	W06597	Hypersensitive respon	3.41e+03	790	4	12.9	472 39	W89559	Triticum sp. cysteine	3.41e+03
718	4	12.9	341 32	W61758	B. burgdorferi antige	3.41e+03	791	4	12.9	473 38	W89835	Protein encoded by cl	3.41e+03
719	4	12.9	344 39	W87641	A hypersensitive resp	3.41e+03	792	4	12.9	474 3	P50587	A.nidulans ribulose-b	3.41e+03
720	4	12.9	344 32	R93797	Protein LP40, a varia	3.41e+03	793	4	12.9	475 16	R77341	Human sonic hedgehog	3.41e+03
721	4	12.9	345 39	W73643	Human placenta short	3.41e+03	794	4	12.9	478 25	W22154	ApXID protein.	3.41e+03
722	4	12.9	347 38	W90057	Rat GlcAT-P protein.	3.41e+03	795	4	12.9	483 38	W82315	Chimeric CD19/ctCR pr	3.41e+03
723	4	12.9	355 30	W40382	S. glaucescens acbC p	3.41e+03	796	4	12.9	489 13	R76697	Mouse MDM2 protein.	3.41e+03
724	4	12.9	357 21	W4773	Nitrile hydratase gen	3.41e+03	797	4	12.9	496 39	W89614	Aspergillus oryzae di	3.41e+03
725	4	12.9	359 39	W73544	Mouse IkappaB-beta pr	3.41e+03	798	4	12.9	496 39	W89597	Aspergillus oryzae am	3.41e+03
726	4	12.9	361 39	W73642	Rat adenosine kinase	3.41e+03	799	4	12.9	496 39	W89586	Aspergillus oryzae am	3.41e+03
727	4	12.9	362 39	W73644	Human placenta long f	3.41e+03	800	4	12.9	497 39	W89786	Staphylococcus aureus	3.41e+03
728	4	12.9	364 39	W95356	IL-1 alpha propiece-E	3.41e+03	801	4	12.9	501 9	R50043	aMEF2.	3.41e+03
729	4	12.9	364 33	W82047	Streptococcus pneumon	3.41e+03	802	4	12.9	504 39	W89391	Human trabecular mesh	3.41e+03
730	4	12.9	375 31	W37784	Plasmodium vivax circ	3.41e+03	803	4	12.9	504 38	W73500	Trabecular meshwork i	3.41e+03
731	4	12.9	378 6	R30609	Subtilisin useful in	3.41e+03	804	4	12.9	505 33	W57851	C. felis esterase, nr	3.41e+03
732	4	12.9	379 28	W31600	Alpha-actinin actin-b	3.41e+03	805	4	12.9	506 39	W83126	PrIII48 lysine speci	3.41e+03
733	4	12.9	379 38	W85036	S. pneumoniae Glnu OR	3.41e+03	806	4	12.9	506 39	W87626	Dirofilaria immitis a	3.41e+03
734	4	12.9	380 39	W89183	Pulmonary surfactant	3.41e+03	807	4	12.9	509 9	R47583	NADH oxidase.	3.41e+03
735	4	12.9	381 3	R14445	Deduced from Mycobact	3.41e+03	808	4	12.9	510 8	R41195	Yeast delta 9 desatur	3.41e+03
736	4	12.9	382 7	R36677	Yeast SMT enzyme.	3.41e+03	809	4	12.9	512 39	W87797	Protease encoded by c	3.41e+03
737	4	12.9	383 39	W70836	Rat-edg, G-protein co	3.41e+03	810	4	12.9	513 39	W87798	Aminopeptidase Slpe e	3.41e+03
738	4	12.9	383 39	W87791	Bordetella pertussis	3.41e+03	811	4	12.9	518 33	W61387	Schizopphyllum commune	3.41e+03
739	4	12.9	384 24	W32378	Homology vector 443-8	3.41e+03	812	4	12.9	522 25	W32222	Avian 3-hydroxy-2-met	3.41e+03
740	4	12.9	387 21	W10726	Hamster beta-2-adren	3.41e+03	813	4	12.9	526 13	R17135	Heliothis virescens r	3.41e+03
741	4	12.9	390 1	P90550	Lipase I from Rhizopu	3.41e+03	814	4	12.9	530 33	W57877	C. felis esterase pro	3.41e+03
742	4	12.9	392 6	R28181	Wild type R. niveus l	3.41e+03	815	4	12.9	532 38	W83141	Chimeric receptor con	3.41e+03
743	4	12.9	392 18	R99482	GDP-Fuc:[beta-D-Gal(1	3.41e+03	816	4	12.9	533 9	R44617	Human retinol binding	3.41e+03
744	4	12.9	394 3	R13750	M. tuberculosis antig	3.41e+03	817	4	12.9	536 32	W46311	AAV4 Rep 68 protein.	3.41e+03
745	4	12.9	400 39	W73764	M. tuberculosis antig	3.41e+03	818	4	12.9	539 39	W87799	Aminopeptidase SlpD e	3.41e+03
746	4	12.9	400 39	W73654	M. tuberculosis antig	3.41e+03	819	4	12.9	544 38	W85028	CDK2-green fluorescen	3.41e+03

820	4	12.9	544 38	W85029	Green fluorescent pro	3.41e+03	833	4	12.9	746 38	W84156	Human desaturase enzy	3.41e+03
821	4	12.9	550 33	W57864	C. felis esterase, nF	3.41e+03	834	4	12.9	752 18	R97834	Kaposi's sarcoma asso	3.41e+03
822	4	12.9	550 33	W57864	C. felis esterase, nF	3.41e+03	835	4	12.9	753 39	W95514	Amino acid sequence o	3.41e+03
823	4	12.9	554 39	W89428	S. pneumoniae histidi	3.41e+03	836	4	12.9	753 39	W83927	Human T85 protein.	3.41e+03
824	4	12.9	559 39	W89803	Staphylococcus aureus	3.41e+03	837	4	12.9	757 6	R30497	N-terminal of LH rece	3.41e+03
825	4	12.9	560 32	P60516	Tox228 diptheria toxi	3.41e+03	838	4	12.9	763 2	R11341	Human Thyroid Stimula	3.41e+03
826	4	12.9	562 38	W83285	Human h-RAB	3.41e+03	839	4	12.9	774 38	W87548	A Tfu DNA polymerase	3.41e+03
827	4	12.9	563 2	R10330	Gene product with lip	3.41e+03	890	4	12.9	777 39	W85601	Hexosaminidase enzyme	3.41e+03
828	4	12.9	568 38	W89828	Protein encoded by cl	3.41e+03	901	4	12.9	797 38	W85042	PKB-green fluorescent	3.41e+03
829	4	12.9	568 13	R73384	Natural resistance-as	3.41e+03	902	4	12.9	797 38	W85043	NFKappaB p65 subunit-	3.41e+03
830	4	12.9	570 39	W88278	Gracilariaopsis lemane	3.41e+03	903	4	12.9	798 38	W85025	Smad4-green fluoresce	3.41e+03
831	4	12.9	573 21	R13728	Delta-1 dehydrogenase	3.41e+03	904	4	12.9	806 38	W85013	Smad4-green fluoresce	3.41e+03
832	4	12.9	573 31	W01657	Human heat shock prot	3.41e+03	905	4	12.9	807 39	W89247	Human PTP04	3.41e+03
833	4	12.9	573 29	W47414	Wild type human L-asp	3.41e+03	906	4	12.9	832 30	W40052	Human F300/CBP-associ	3.41e+03
834	4	12.9	575 9	R47175	Sequence of polypepti	3.41e+03	907	4	12.9	836 38	W85017	Grk5-green fluorescen	3.41e+03
835	4	12.9	583 18	R97230	Stem cell marker HCAP	3.41e+03	908	4	12.9	842 38	W85008	Grk5-green fluorescen	3.41e+03
836	4	12.9	584 24	W12661	Orange pectin methyle	3.41e+03	909	4	12.9	842 30	W56698	Neuronal migration pr	3.41e+03
837	4	12.9	587 25	W33274	S. fradiae tylosin bi	3.41e+03	910	4	12.9	842 12	R63570	Feedback inhibition r	3.41e+03
838	4	12.9	589 2	R10681	Polyhydroxybutyrate p	3.41e+03	911	4	12.9	842 12	R63568	Feedback inhibition r	3.41e+03
839	4	12.9	590 39	W87628	Dirofilaria immitis a	3.41e+03	912	4	12.9	843 38	W85030	Green fluorescent pro	3.41e+03
840	4	12.9	590 39	W87629	Dirofilaria immitis a	3.41e+03	913	4	12.9	844 29	W40119	Human GABA-BR1b recep	3.41e+03
841	4	12.9	590 1	P96204	Human muscarinic acet	3.41e+03	914	4	12.9	847 39	W85603	Hexosaminidase enzyme	3.41e+03
842	4	12.9	591 38	W73316	Parathyroid hormone r	3.41e+03	915	4	12.9	853 38	W85031	Protein tyrosine phos	3.41e+03
843	4	12.9	593 38	W73317	Human Parathyroid hor	3.41e+03	916	4	12.9	864 39	W30576	Pyruvate formate lyas	3.41e+03
844	4	12.9	595 38	W85021	mPKAc-green fluoresce	3.41e+03	917	4	12.9	865 38	W85027	Zap70-green fluoresce	3.41e+03
845	4	12.9	604 38	W85016	Erk2-green fluorescen	3.41e+03	918	4	12.9	871 38	W85026	Green fluorescent pro	3.41e+03
846	4	12.9	605 38	W85007	Erk2-green fluorescen	3.41e+03	919	4	12.9	873 19	W02212	Human VLDL receptor.	3.41e+03
847	4	12.9	606 38	W85019	p38-green fluorescen	3.41e+03	920	4	12.9	886 39	W89787	Staphylococcus aureus	3.41e+03
848	4	12.9	607 38	W85010	p38-green fluorescen	3.41e+03	921	4	12.9	890 39	W73534	MEK4 protein.	3.41e+03
849	4	12.9	607 17	R91016	PRSPG gluc fusion pro	3.41e+03	922	4	12.9	892 28	W34260	Rat ob receptor isofo	3.41e+03
850	4	12.9	613 8	R40106	Pseudomonas exotoxin	3.41e+03	923	4	12.9	906 1	P91047	gB envelope protein o	3.41e+03
851	4	12.9	613 8	R40108	Pseudomonas exotoxin	3.41e+03	924	4	12.9	908 38	W85587	Human N-methyl-D-aspa	3.41e+03
852	4	12.9	620 33	W50907	Alteromonas carragee	3.41e+03	925	4	12.9	916 38	W85023	Amino acid sequence o	3.41e+03
853	4	12.9	623 38	W88707	Secreted protein enco	3.41e+03	926	4	12.9	917 38	W73512	Rat ICAM-4 protein se	3.41e+03
854	4	12.9	623 32	W46312	AAV4 Rep 78 protein.	3.41e+03	927	4	12.9	924 38	W73511	Human neuronal ICAM-4	3.41e+03
855	4	12.9	624 38	W85015	Erk1-green fluorescen	3.41e+03	928	4	12.9	924 8	R42385	Recombinant leukotoxi	3.41e+03
856	4	12.9	625 39	W67882	Human secreted protei	3.41e+03	929	4	12.9	926 3	R14482	LXN352.	3.41e+03
857	4	12.9	630 38	W85018	Jnk1-green fluorescen	3.41e+03	930	4	12.9	929 12	R66056	Human NMDA receptor s	3.41e+03
858	4	12.9	631 38	W85006	Erk1-green fluorescen	3.41e+03	931	4	12.9	929 38	W87501	N-methyl-D-aspartate	3.41e+03
859	4	12.9	633 38	W85009	Jnk1-green fluorescen	3.41e+03	932	4	12.9	930 39	W75426	Human aggrecan degrad	3.41e+03
860	4	12.9	635 38	W75410	Fusion protein PD30 c	3.41e+03	933	4	12.9	931 39	W89596	Saccharomyces cerevis	3.41e+03
861	4	12.9	635 38	W75412	Fusion protein PD33 c	3.41e+03	934	4	12.9	931 15	R79009	Human DNA repair prot	3.41e+03
862	4	12.9	635 38	W75409	Fusion protein PD31 c	3.41e+03	935	4	12.9	932 38	W85040	PKG-green fluorescent	3.41e+03
863	4	12.9	635 38	W85034	Green fluorescent pro	3.41e+03	936	4	12.9	933 39	W89200	Human phosphatidylicho	3.41e+03
864	4	12.9	635 25	W25755	PHA depolymerase.	3.41e+03	937	4	12.9	933 38	W85039	Green fluorescent pro	3.41e+03
865	4	12.9	640 24	W10065	Human heat shock prot	3.41e+03	938	4	12.9	936 39	W89801	Staphylococcus aureus	3.41e+03
866	4	12.9	646 24	W27178	Nephila clavipes spid	3.41e+03	939	4	12.9	936 15	R80251	Polymer SELP0.	3.41e+03
867	4	12.9	647 39	W73573	GST-Msp I methylase f	3.41e+03	940	4	12.9	945 39	W73624	Human secreted protei	3.41e+03
868	4	12.9	649 38	W81977	Ehrlichia sp. B3 prot	3.41e+03	941	4	12.9	955 8	R42237	Human p50 protein DNA	3.41e+03
869	4	12.9	668 7	R32457	PE having M1 residues	3.41e+03	942	4	12.9	955 8	R42238	Human p50 protein DNA	3.41e+03
870	4	12.9	670 7	R32468	BSPEN1c5aa fragment.	3.41e+03	943	4	12.9	968 38	W85011	p85alpha-green floure	3.41e+03
871	4	12.9	672 39	W73536	MEK1 protein.	3.41e+03	944	4	12.9	970 38	W85020	p85alpha-green floure	3.41e+03
872	4	12.9	674 28	W31345	Cryptosporidium parvu	3.41e+03	945	4	12.9	970 14	R72458	Porphyromonas gingiva	3.41e+03
873	4	12.9	677 38	W75411	Fusion protein PD34 c	3.41e+03	946	4	12.9	975 33	W89801	Amino acid sequence o	3.41e+03
874	4	12.9	682 21	W17606	Thermoanaerobacter CG	3.41e+03	947	4	12.9	976 38	W85581	Human N-methyl-D-aspa	3.41e+03
875	4	12.9	683 21	W17614	Thermoanaerobacter CG	3.41e+03	948	4	12.9	990 39	W90075	P. vulgaris chondroit	3.41e+03
876	4	12.9	683 21	W17614	Thermoanaerobacter CG	3.41e+03	949	4	12.9	996 38	W85033	IkappaB kinase alpha	3.41e+03
877	4	12.9	691 31	W53320	Staphylococcus aureus	3.41e+03	950	4	12.9	997 38	W85589	Human N-methyl-D-aspa	3.41e+03
878	4	12.9	697 27	R732324	Staphylococcus aureus	3.41e+03	951	4	12.9	997 38	W85032	Human N-methyl-D-aspa	3.41e+03
879	4	12.9	706 39	W81353	Human hrizled-6 pro	3.41e+03	952	4	12.9	1003 6	R29705	pol gene decoded from	3.41e+03
880	4	12.9	711 31	W57321	Topoisomerase I expre	3.41e+03	953	4	12.9	1010 18	R91823	Human immunodeficienc	3.41e+03
881	4	12.9	718 38	W85024	Erk2-green fluorescen	3.41e+03	954	4	12.9	1039 38	W85014	Stat5-green fluoresce	3.41e+03
882	4	12.9	719 39	W84067	Drosophila mitofusin	3.41e+03	955	4	12.9	1045 38	W85050	Stat5-green fluoresce	3.41e+03
883	4	12.9	719 38	W85012	Smad2-green fluoresce	3.41e+03	956	4	12.9	1066 39	W88255	Morchella costata alp	3.41e+03
884	4	12.9	726 38	W85022	PKB-green fluorescen	3.41e+03	957	4	12.9	1070 39	W88256	Morchella vulgaris al	3.41e+03
885	4	12.9	726 3	R15033	KGF receptor.	3.41e+03	958	4	12.9	1086 38	W81839	Human LFA-1 alpha cha	3.41e+03
886	4	12.9	727 38	W85041	Green fluorescent pro	3.41e+03	959	4	12.9	1089 38	W82396	Human UBP protein #2.	3.41e+03
887	4	12.9	731 39	W87767	Human gelsolin.	3.41e+03	960	4	12.9	1091 39	W88254	Gracilariaopsis lemane	3.41e+03
888	4	12.9	732 33	W63715	Rat rsk3 protein.	3.41e+03	961	4	12.9	1092 39	W88257	Gracilariaopsis lemane	3.41e+03
889	4	12.9	740 39	W95398	M. tuberculosis catal	3.41e+03	962	4	12.9	1101 38	W83160	Rat orphan tyrosine k	3.41e+03
890	4	12.9	740 39	W95399	M. tuberculosis catal	3.41e+03	963	4	12.9	1115 3	R13457	Duffy receptor.	3.41e+03
891	4	12.9	740 39	W83066	Pyruvate formate lyas	3.41e+03	964	4	12.9	1162 26	W23398	Rat ob receptor (wild	3.41e+03
892	4	12.9	746 38	W85135	A desaturase enzyme e	3.41e+03	965	4	12.9	1171 38	W85037	Green fluorescent pro	3.41e+03

```
966 4 12.9 1181 38 W85038 NFAR1-green fluoresce 3.41e+03
967 4 12.9 1189 33 W61338 A mutated CryIC prote 3.41e+03
968 4 12.9 1195 1 R05530 High density lipoprot 3.41e+03
969 4 12.9 1203 39 W81572 Mus dunni endogenous 3.41e+03
970 4 12.9 1237 39 W89347 Chicken transmembrane 3.41e+03
971 4 12.9 1247 39 W73535 MEK5 protein. 3.41e+03
972 4 12.9 1262 25 W13505 B. bronchiseptica ade 3.41e+03
973 4 12.9 1265 38 W68594 Human methionine synt 3.41e+03
974 4 12.9 1292 1 R05531 High density lipoprot 3.41e+03
975 4 12.9 1358 26 W35743 Recombinant human res 3.41e+03
976 4 12.9 1388 31 W56475 Protein with rho prot 3.41e+03
977 4 12.9 1405 31 W56101 Enhanced green fluore 3.41e+03
978 4 12.9 1447 17 W75375 Human patched protein 3.41e+03
979 4 12.9 1451 38 W83310 LRP5 protein from iso 3.41e+03
980 4 12.9 1483 22 W19700 ATM mutant 4443del175 3.41e+03
981 4 12.9 1523 38 W87554 A Tfu DNA polymerase 3.41e+03
982 4 12.9 1591 38 W83308 Mature LRP5 protein. 3.41e+03
983 4 12.9 1614 38 W83312 Mouse lrp5 protein. 3.41e+03
984 4 12.9 1615 38 W83309 LRP5 protein from the 3.41e+03
985 4 12.9 1639 38 W83311 LRP5 isoform 3 protei 3.41e+03
986 4 12.9 1670 18 R94616 DNA polymerase. 3.41e+03
987 4 12.9 1879 9 R42455 Enzyme involved in ei 3.41e+03
988 4 12.9 2042 32 W63319 Haemophilus paragall 3.41e+03
989 4 12.9 2055 31 W53522 Amino acid sequence o 3.41e+03
990 4 12.9 2107 31 W53521 Amino acid sequence o 3.41e+03
991 4 12.9 2189 1 R05222 Antigen GX5401FL enco 3.41e+03
992 4 12.9 2233 33 W48711 HPV-3 JS isolate wil 3.41e+03
993 4 12.9 2476 39 W67738 Pig p105 zona pelluci 3.41e+03
994 4 12.9 2998 22 W19691 ATM mutant 3403del174 3.41e+03
995 4 12.9 3046 22 W19690 ATM mutant C9140T. 3.41e+03
996 4 12.9 3054 22 W19702 ATM mutant 7278del6. 3.41e+03
997 4 12.9 3587 7 R34713 Bacillus subtilis srf 3.41e+03
998 4 12.9 3910 7 R38470 ALL-1 protein. 3.41e+03
999 4 12.9 4302 28 W28830 Human PKD1 protein. 3.41e+03
1000 4 12.9 4655 29 W43311 Human calcium sensor 3.41e+03
```

## ALIGNMENTS

```
RESULT 1
ID W06548 standard; Protein; 228 AA.
AC W06548;
DT 13-MAR-1997 (first entry)
DE Human colon specific gene CSG5 polypeptide fragment.
KW Colon specific gene; CSG5; colon cancer; metastasis; diagnosis;
KW therapy; antibody; vaccine.
OS Homo sapiens.
PN W09639419-Al.
PD 12-DEC-1996.
PF 06-JUN-1995; U07289.
PR 06-JUN-1995; WO-U07289.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Yu G;
DR N-PSDB; T45884.
PT Human colon specific genes and their expression products - detection
PT of which, in non-colon tissue samples, can be used as indication of
PT colon cancer metastasis
PS Claim 8; Fig 5; 60pp; English.
CC Novel polypeptides (W06545-53) are encoded by cDNA clones (see also
CC T45880-92) corresponding to 13 human colon specific genes,
CC designated CSG1, CSG2, etc., that are primarily expressed in
CC tissues derived from the colon. Recombinant CSG polypeptides can
CC be produced in transformed host cells. They are useful diagnostic
CC markers for colon cancer and for colon cancer metastasis and can
CC also be used to screen for (antagonist cpds. of therapeutic or
CC diagnostic value. Antibodies raised against the colon-specific
CC polypeptides may be used to target colon cancer cells or as part
CC of a colon cancer vaccine.
SQ Sequence 228 AA;
```

Query Match 58.1%; Score 18; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.66e-12;

```
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 VYSRYFTYDNGRYSVK 18
QY 14 VYSRYFTYDNGRYSVK 31
RESULT 2
ID W46879 standard; Protein; 228 AA.
AC W46879;
DT 22-JUN-1998 (first entry)
DE Protein sequence encoded by a colon-specific gene.
KW Colon-specific gene; probe; detection; expression; human;
KW diagnostic assay; colon cancer; antibody; screening.
OS Homo sapiens.
PN US5733748-A.
PD 31-MAR-1998.
PF 06-JUN-1995; 469667.
PR 06-JUN-1995; US-469667.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen C, Yu G;
DR WPI; 98-229823/20.
DR N-PSDB; V16672.
DE Colon-specific nucleic acids - useful as probes for detecting colon
DE cancer micrometastases
PS Claim 1; Fig 5; 51pp; English.
CC W46876-80 and W4682-85 represent proteins encoded by colon-specific
CC genes. The polynucleotides encoding these proteins can be used
CC as probes to detect expression of the corresponding human genes,
CC e.g. in diagnostic assays for detecting micrometastases of colon cancer.
CC Recombinant cells containing the polynucleotides can be used to
CC produce the proteins, in order that antibodies can be raised and
CC used in further screening or diagnostics.
SQ Sequence 228 AA;
```

Query Match 58.1%; Score 18; DB 29; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.66e-12;

```
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 VYSRYFTYDNGRYSVK 18
QY 14 VYSRYFTYDNGRYSVK 31
RESULT 3
ID W41166 standard; Protein; 345 AA.
AC W41166;
DT 30-JUL-1998 (first entry)
DE Metal-regulated transporter polypeptide IRT2.
KW MRT; transgenic plant; pollution; metal-deficiency.
OS Arabidopsis thaliana.
PN W09745000-Al.
PD 04-DEC-1997.
PF 27-NOV-1996; U19065.
PR 11-OCT-1996; CA-187728.
PR 29-MAY-1996; US-018578.
PA (DART-) DARTMOUTH COLLEGE.
PA (MINU) UNIV MINNESOTA.
PI Elide DJ, Guerinot ML;
DR WPI; 98-032230/03.
DR N-PSDB; V12767.
PT Novel A. thaliana metal-regulated transporter proteins and related
PT DNA - for generating transgenic plants useful for removing
PT pollutants from soil or providing nutrient to patients suffering
PT from metal-deficiency disorders
PS Claim 33; Fig 23; 144pp; English.
CC The metal-regulated transporter (MRT) polypeptides (W41162-W41166) are
CC capable of transporting metals such as Fe(II), Pb and Zn. The peptides
CC can be altered to generate transgenic plants in which the expression of MRT
CC can be used to remove pollutants from the soil or to treat a disorder associated with metal-deficiency. The
CC polypeptides can also be used to identify agents which inhibit or
CC stimulate MRT or MRT expression.
```



```

PF 05-FEB-1996; DK0057.
PR 03-FEB-1995; DK-000128.
PR 23-OCT-1995; DK-001192.
PR 10-NOV-1995; DK-001256.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgard-frantzen H, Borchert TV, Svendsen A;
DR WPI: 96-371424/37.
PT Alpha-amylase variants and methods of production - have altered
PT properties such as calcium dependency, substrate binding and
PT stability
PS Disclosure; Page 86; 171pp; English.
CC The present sequence is the mature Bacillus amyloliquefaciens alpha-
CC amylase (A). Variants of parent termamyl-like alpha-amylases (and methods
CC of constructing them) are claimed. Examples of variants are featured
CC above. At least one of the amino acids of the parent, which is present in
CC a fragment identified above, is/are deleted or replaced with one or more
CC residues, which is/are present in a fragment corresponding to a fragment
CC of W14500 (Aspergillus oryzae fungamyl alpha-amylase). The variants have
CC altered properties such as calcium dependency, substrate binding and
CC stability. Also one or more proline or cysteine residues in the variant
CC is modified or replaced with a non-proline or non-cysteine residue such
CC as alanine. The variants can be used for (dish)washing, as detergent
CC additives or for fabric desizing or starch liquefaction. They can also
CC be used for the production of sweeteners and ethanol from starch. See
CC also W14498 and W14500.
SQ Sequence 480 AA;

Query Match 19.4%; Score 6; DB 21; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.44e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 agadat 110
|||||
Qy 4 AGADAT 9

RESULT 6
ID R78268 standard; Protein; 480 AA.
AC R78268;
DE 17-JAN-1996 (first entry)
DE Bacillus amyloliquefaciens alpha amylase (mature protein).
KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
KW starch; thermostable; methionine; Bacillus licheniformis;
KW Bacillus amyloliquefaciens; Bacillus stearothermophilus.
OS Bacillus amyloliquefaciens.
PN WO9521247-A1.
PD 10-AUG-1995.
PF 05-OCT-1994; DK0371.
PR 02-FEB-1994; DK-000141.
PA (NOVO ) NOVO-NORDISK AS.
PI Marcher D, Nilsson TE, Pedersen HH, Toft AH;
DR WPI: 95-283767/37.
DR N-PSDB: Q95032.
PT Use of an oxidation stable alpha-amylase - for simultaneous desizing
PT and bleaching or scouring of fabrics contg. starch or starch derivs.
PS Claim 7; Page 24; 37pp; English.
CC Oxidation stable alpha amylases can be used for the simultaneous
CC desizing and bleaching or scouring of a fabric comprising starch or
CC starch derivatives. They exhibit a better heat stability,
CC especially in the presence of oxidising agents. They are obtained
CC from a parent alpha amylase by replacing one or more methionine
CC residues with any amino acid different from Cys or Met, preferably
CC Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is
CC derived from a Bacillus species. This sequence is the wild type
CC (unmodified) alpha amylase.
SQ Sequence 480 AA;

Query Match 19.4%; Score 6; DB 14; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.44e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 agadat 110
|||||
Qy 4 AGADAT 9

RESULT 7
ID W31405 standard; Protein; 483 AA.
AC W31405;
DE 11-MAY-1998 (first entry)
DE Bacillus amyloliquefaciens Termamyl-like alpha-amylase.
KW Termamyl; alpha-amylase; enzyme engineering; protein engineering;
KW starch; liquefaction; saccharification; sweetener;
KW textile desizing; detergent additive; ss.
OS Bacillus amyloliquefaciens.
PN WO9741213-A1.
PD 06-NOV-1997.
PF 30-APR-1997; DK0197.
PR 08-NOV-1996; DK-001263.
PR 30-APR-1996; DK-000515.
PR 28-JUN-1996; DK-000712.
PR 11-JUL-1996; DK-000775.
PA (NOVO ) NOVO-NORDISK AS.
PI Bisgard-Frantzen H, Borchert TV, Svendsen A;
DR WPI: 97-549718/50.
DR N-PSDB: V02472.
PT Termamyl-like alpha-amylase variants with improved properties - e.g.
PT increased stability at low pH and low calcium, useful as detergent
PT additives and in industrial starch processing e.g. liquefaction
PS Disclosure; Page 84; 101pp; English.
CC This protein comprises the Termamyl-like alpha-amylase of Bacillus
CC amyloliquefaciens. The invention relates to novel variants of
CC Termamyl-like alpha-amylases that have alpha-amylase activity and
CC exhibit an alteration in at least one property selected from:
CC substrate specificity; binding or cleavage pattern; thermal
CC stability; pH/activity or pH/stability profile; stability towards
CC oxidation; Ca2+ dependency and specific activity. The variant has
CC one or more mutations from those listed in the specification in
CC relation to Bacillus licheniformis Termamyl (see W31404). Also
CC claimed are constructs comprising DNA encoding the variant (see
CC V02471-73), and recombinant expression vectors and transformed
CC cells containing the DNA. The Termamyl-like alpha-amylase variant
CC is useful as a detergent additive and can also be used in
CC industrial starch processing e.g. liquefaction (claimed) or
CC saccharification to produce sweeteners, and in textile desizing
CC (claimed).
SQ Sequence 483 AA;

Query Match 19.4%; Score 6; DB 28; Length 483;
Best Local Similarity 100.0%; Pred. No. 5.44e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 agadat 110
|||||
Qy 4 AGADAT 9

RESULT 8
ID R72448 standard; Protein; 483 AA.
AC R72448;
DE 01-DEC-1995 (first entry)
DE Bacillus amyloliquefaciens alpha amylase (mature protein).
KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
KW thermostable.
OS Bacillus amyloliquefaciens.
PN WO9510603-A.
PD 20-APR-1995.
PF 05-OCT-1994; DK0370.
PR 08-OCT-1993; DK-001133.
PR 02-FEB-1994; DK-000140.
PA (NOVO ) NOVO-NORDISK AS.
PI Svendsen A, Thellertsen M;
PI Van der zee P, Bisgardfrantzen H, Borchert T;
DR WPI: 95-161790/21.

```

DR N-PSDB; Q88067.  
 PT New Bacillus derived alpha-amylase variants - having amino acid  
 PT modifications to improve washing and/or dishwashing performance  
 PS Claim 34; Page 11-12; 105pp; English.  
 CC Variant alpha amylase enzymes which have improved washing and/or  
 CC as detergent additives. The enzymes have one or more amino acid  
 CC residues added, deleted or substituted. The variants can also be  
 CC used for textile desizing prior to scouring, bleaching and dyeing.  
 CC The variants have improved thermostability, acid/alkaline stability;  
 CC low temperature optimum; pH optimum; higher hydrolysis velocity and  
 CC improved tolerance to other composition constituents, e.g. oxidation  
 CC agents.  
 SQ Sequence 483 AA;

Query Match 19.4%; Score 6; DB 14; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 5.44e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 105 agadat 110  
 |||||  
 QY 4 AGADAT 9

## RESULT 9

ID R10378 standard; Protein; 514 AA.

AC R10378;  
 DE 05-APR-1991 (first entry)  
 DE Mutant alpha-amylase gene product.  
 KW PUCAM4.

OS Bacillus amyloliquefaciens.

FH Key Location/Qualifiers

FT protein 32..514

FT misc\_difference 154 /label= Mature protein

FT /label= Arg-Cys mutation site

PN EP-409299-A.

PD 23-JAN-1991.

PF 27-JUN-1990; 201704.

PR 29-JUN-1989; EP-201732.

PR 27-JUN-1990; EP-201704.

PA (KONN ) GIST-BROCADES NV.

PI Van Ejik JH, Quax WJ, Sanders JP;

DR WPI; 91-023909/04.

DR N-PSDB; Q10322.

PT Mutant enzyme having reduced stability - comprising modified

PT alpha-amylase which can be used in baking to improve loaf vol.

PT and crumb softness without over-dextrinisation

PS Claim 18; Fig 3; 26pp; English.

CC The wild type amylase sequence differs at Arg123-Cys, and optionally

CC at positions 113-4, 116, 123, 163-4, 166, 238, 316, 322, 345, 349,

CC 356, 386, 394 or 398. The mutant for is less stable under industrial

CC conditions, and may be used as active ingredient in bread making,

CC improving crumb softness and loaf volume without over dextrinisation

CC of starch.

CC Sequence 514 AA;

Query Match 19.4%; Score 6; DB 2; Length 514;

Best Local Similarity 100.0%; Pred. No. 5.44e+01;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 136 agadat 141

|||||

QY 4 AGADAT 9

## RESULT 10

ID R10379 standard; Protein; 514 AA.

AC R10379;

DT 05-APR-1991 (first entry)

DE Wild type alpha-amylase gene product.

KW PUCAM4.

OS Bacillus amyloliquefaciens.

FH Key Location/Qualifiers

FT protein 32..514  
 FT misc\_difference 154 /label= Mature protein  
 FT /label= Arg-Cys mutation site  
 PN EP-409299-A.  
 PD 23-JAN-1991.  
 PF 27-JUN-1990; 201704.  
 PR 29-JUN-1989; EP-201732.  
 PR 27-JUN-1990; EP-201704.  
 PA (KONN ) GIST-BROCADES NV.  
 PI Van Ejik JH, Quax WJ, Sanders JP;  
 DR WPI; 91-023909/04.  
 DR N-PSDB; Q10322.  
 PT Mutant enzyme having reduced stability - comprising modified  
 PT alpha-amylase which can be used in baking to improve loaf vol.  
 PT and crumb softness without over-dextrinisation  
 PS Claim 18; Fig 3; 26pp; English.  
 CC The wild type amylase sequence differs at Arg123-Cys, and optionally  
 CC at positions 113-4, 116, 123, 163-4, 166, 238, 316, 322, 345, 349,  
 CC 356, 386, 394 or 398. The mutant for is less stable under industrial  
 CC conditions, and may be used as active ingredient in bread making,  
 CC improving crumb softness and loaf volume without over dextrinisation  
 CC of starch.  
 SQ Sequence 514 AA;

Query Match 19.4%; Score 6; DB 2; Length 514;

Best Local Similarity 100.0%; Pred. No. 5.44e+01;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 136 agadat 141

|||||

QY 4 AGADAT 9

## RESULT 11

ID W39743 standard; Protein; 520 AA.

AC W39743;

DT 20-MAY-1998 (first entry)

DE B. amyloliquefaciens alpha amylase protein.

KW Alpha-amylase; calcium binding; starch liquefaction; detergent;

KW baking aid; textile industry.

OS Bacillus amyloliquefaciens.

PN W09743424-Al.

PD 20-NOV-1997.

PF 06-MAY-1997; U07609.

PR 14-MAY-1996; US-645971.

PA (GENV ) GENECOR INT INC.

PI Bott RR, Shaw A;

DR WPI; 98-008893/01.

DR New modified alpha-amylase enzymes - having altered calcium binding

PT properties to alter e.g. high or low pH activity, thermostability or

PT oxidative stability

PS Disclosure; Fig 5A; 31pp; English.

CC This sequence represents an alpha-amylase isolated from Bacillus

CC amyloliquefaciens. This sequence is used in the analysis of novel

CC alpha-amylases comprising of an A domain, a C domain and a calcium

CC binding site associated with the A domain and the C domain. Ligand

CC residues in the A and/or C domains allow the modification of the

CC alpha-amylase which alter the characteristics of the calcium binding

CC site and thereby alter performance. This modified enzyme can be used

CC for e.g. starch liquefaction, in laundry or dishwashing detergent

CC compositions, as baking aids or in textile desizing. The alpha-amylase

CC variants can have increased activity at low pH and high temperatures,

CC increased high pH and oxidative stability and improved stability in

CC the absence or low concentrations of calcium ion.

CC Sequence 520 AA;

Query Match 19.4%; Score 6; DB 28; Length 520;

Best Local Similarity 100.0%; Pred. No. 5.44e+01;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 142 agadat 147

|||||

QY 4 AGADAT 9

## RESULT 12

ID R10100 standard; Protein; 615 AA.  
 AC R10100:  
 DT 14-MAR-1991 (first entry)  
 DE E.tenella protein Et1A1.  
 KW Lambda gt10Et1A1; poultry coccidiosis; vaccine.  
 OS Eimeria tenella.  
 PN EP-403698-A.  
 PD 27-DEC-1990.  
 PF 20-DEC-1989; 203262.  
 PR 21-JUN-1989; ZA-004726.  
 PA (ALKU ) AKZO NV.  
 PI Vermeulen A., Kok JJ, Dijkema R, Van Den Boogaart P;  
 DR WPI: 91-000856/01.  
 DQ Q-PSDB: Q10156.  
 PT New nucleic acid encoding Eimeria tenella protein - and derived  
 PT vectors, host cells, polypeptide(s) and antibodies, useful in  
 PT vaccines against poultry coccidiosis.  
 PS Claim 1: Fig 2: 1lpp; English.  
 CC Peptide product may be produced from a transformed expression system,  
 CC and used in vaccination against E.tenella infection. Abs raised to  
 CC the peptide may be used in treatment and diagnosis of infected fowl.  
 CC Sequence 615 AA;

Query Match 19.4%; Score 6; DB 2; Length 615;  
 Best Local Similarity 100.0%; Pred.No. 5.44e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 166 gagada 171

QY 3 GAGADA 8

## RESULT 13

ID W72748 standard; Protein; 674 AA.  
 AC W72748:  
 DT 01-FEB-1999 (first entry)  
 DE Human P-dlg protein.  
 KW Human; P-dlg; dlg family; detection; carcinostatic mechanism;  
 KW anticancer; tumour suppressor; embryogenesis; regulation.  
 OS Homo sapiens.  
 FH Key  
 FT Location/Qualifiers  
 FT Misc\_difference 562  
 FT /label= unknown  
 FT /note= "encoded by WAG"  
 FT W09846745-A1.  
 PD 22-OCT-1998.  
 PF 08-APR-1998; J01611.  
 PR 14-APR-1997; JP-111846.  
 PA (SUWE ) SUMITOMO ELECTRIC IND CO.  
 PI Nakamura H, Nakata M, Saya H;  
 DR WPI: 98-568727/48.  
 DQ N-PSDB: V67191.  
 PT New human dlg family protein - used for, e.g. investigation of  
 PT mechanisms of tumour suppression and for development of anticancer  
 PT agents  
 PS claim 1; Page 29-32; 55pp; Japanese.  
 CC The present sequence is a human protein belonging to the dlg family  
 CC (P-dlg). P-dlg is recognised by an antibody binding to the epitope:  
 CC KEQDPDYLRLDKVQRHSKE. The dlg family of proteins is involved in  
 CC regulation of embryogenesis and in tumour suppression. P-dlg, and its  
 CC associated polynucleotides and antibodies are useful in investigating  
 CC mechanisms of tumour suppression and in the design and screening of  
 CC potential anticancer agents.  
 CC Sequence 674 AA;

Query Match 19.4%; Score 6; DB 37; Length 674;  
 Best Local Similarity 100.0%; Pred.No. 5.44e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 635 ysyft 640  
 QY 15 YSYFT 20

## RESULT 14

ID R89275 standard; Protein; 729 AA.  
 AC R89275:  
 DT 25-JUL-1996 (first entry)  
 DE Yeast coagulation protein FLO8.  
 KW Saccharomyces cerevisiae; coagulation; FLO8; mutant; wild type; wine;  
 KW beer; yeast; flavour; alcoholic beverage.  
 OS Saccharomyces cerevisiae.  
 PN J08000270-A.  
 PD 09-JAN-1996.  
 PF 23-JUN-1994; 141520.  
 PR 23-JUN-1994; JP-141520.  
 PA (KIRI ) KIRIN BREWERY KK.  
 DR WPI: 96-091654/10.  
 DQ N-PSDB: Q99257.  
 PT Yeast coagulating gene FLO8 and prepn. of yeast with increased  
 PT coagulation - also prepn. of yeast with decreased coagulation, both  
 PT useful in wine making and brewing industries  
 PS Claim 1; Page 13-16; 18pp; Japanese.  
 CC This is the amino acid sequence of the Saccharomyces cerevisiae  
 CC coagulation protein FLO8. The corresp. gene was isolated by screening  
 CC a non-coagulating yeast strain (YPH500) transformed with a library  
 CC generated from DNA from a coagulating yeast strain (ATCC-60715).  
 CC The full length gene can be used to generate mutants having reduced  
 CC or increased coagulation activity. The strain carrying the wild type  
 CC or mutant FLO8 genes can be used in wine and beer making, where the  
 CC coagulating activity of the yeast can affect the flavour of the  
 CC alcoholic beverage.  
 CC Sequence 729 AA;

Query Match 19.4%; Score 6; DB 17; Length 729;  
 Best Local Similarity 100.0%; Pred.No. 5.44e+01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 435 adatkd 440

QY 6 ADATKD 11

## RESULT 15

ID W33624 standard; Protein; 996 AA.  
 AC W33624:  
 DT 21-MAY-1998 (first entry)  
 DE Eimeria tenella 45 kDa sporozoite surface antigen.  
 KW Coccidiosis; vaccine; poultry; protozoan; parasite; antigen;  
 KW sporozoite.  
 OS Eimeria tenella.  
 PN US5661015-A.  
 PD 26-AUG-1997.  
 PF 03-JUN-1988; 202721.  
 PR 20-DEC-1991; US-812349.  
 PR 03-JUN-1988; US-202721.  
 PA (HOFF ) HOFFMANN LA ROCHE INC.  
 PI Altenburger W, Binger M, Chizzonite RA, Kramer RA,  
 PI Lomedico PT, McAndrew SJ;  
 DR WPI: 97-434379/40.  
 DQ N-PSDB: T93597.  
 PT New DNA from Eimeria tenella and related immunogenic polypeptides -  
 PT useful in vaccines to protect poultry against coccidiosis  
 PS Claim 2; Fig 3A-B; 72pp; English.  
 CC This 45 kDa protein is recognised by monoclonal antibody 7B2  
 CC (ATCC HB 9712). This antibody also specifically reacts with an  
 CC Eimeria tenella 200 kDa surface antigen that is present in the  
 CC sporozoite developmental stage. The amino acid sequence was  
 CC deduced from a cDNA clone (see T93597) obtained from a cDNA  
 CC library by immunological screening with monoclonal antibodies  
 CC raised against Eimeria antigens. It differs from the sequence  
 CC given in W33621 by deletion of 2 N-terminal amino acids that



CC are encoded by nucleotide linker sequences. The invention provides  
CC DNA sequences (see T93593-98) coding for Eimeria surface antigens  
CC (see W31582-84 and W33621-26), recombinant vectors containing such  
CC DNA sequences, transformed microorganisms containing such vectors,  
CC and methods for producing the antigens using the transformed  
CC microorganisms. Methods are also provided for protecting poultry  
CC against coccidiosis using the Eimeria surface antigens. The  
CC surface antigens are administered either as purified proteins or in  
CC the form of DNA encoding the proteins in a viral vector such as a  
CC vaccinia virus. The vaccines may produce antibodies that are  
CC cross-reactive with other Eimeria species.  
SQ Sequence 996 AA:

Query Match 19.4%; Score 6; DB 28; Length 996;  
Best Local Similarity 100.0%; Pred.No. 5.44e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 507 gagada 512  
Qy 3 GAGADA 8  
|||||

Search completed: Sat Aug 28 15:05:06 1999  
Job time : 70 secs.

**This Page Blank (uspto)**



97	140	1	HBFGRE	hemoglobin beta chain	2.73e+02	170	245	2	H71429	hypothetical protein	2.73e+02
98	141	1	UDHUP2	cystatin SN precursor	2.73e+02	171	245	2	S39481	glutamate--ammonia li	2.73e+02
99	142	2	S04664	T-cell receptor delta	2.73e+02	172	247	2	C69954	hypothetical protein	2.73e+02
100	142	2	S38307	T-cell receptor delta	2.73e+02	173	247	1	D24706	modulation protein no	2.73e+02
101	144	2	S38308	T-cell receptor delta	2.73e+02	174	248	2	A69173	conserved hypothetica	2.73e+02
102	144	2	S36322	T-cell receptor delta	2.73e+02	175	249	2	S66851	probable membrane pro	2.73e+02
103	145	2	S36324	T-cell receptor delta	2.73e+02	176	249	2	H69531	conserved hypothetica	2.73e+02
104	146	1	GGGAA	globin - slug sea har	2.73e+02	177	251	2	C36044	indole-3-glycerol-pho	2.73e+02
105	146	1	GGGDA	globin - sea hare (Do	2.73e+02	178	252	2	C69490	LSU ribosomal protein	2.73e+02
106	146	2	B70560	hypothetical protein	2.73e+02	179	253	2	A64456	hypothetical protein	2.73e+02
107	146	2	A30223	globin - ragged sea h	2.73e+02	180	253	1	VHVTVV	nucleoscapid protein	2.73e+02
108	147	1	GGNW1B	globin, major monomer	2.73e+02	181	254	2	S49329	trypsin-like protein	2.73e+02
109	148	2	G70890	probable repressor -	2.73e+02	182	254	2	D59648	2-keto-3-deoxyglucon	2.73e+02
110	148	2	A33420	globin IV - polychaet	2.73e+02	183	255	2	T03482	iron(iii) dicitrate t	2.73e+02
111	150	2	S36312	T-cell receptor delta	2.73e+02	184	259	1	DEECSP	sorbitol-6-phosphate	2.73e+02
112	150	2	I46624	rearranged T-cell rec	2.73e+02	185	259	2	S14731	archaerhodopsin 2 - H	2.73e+02
113	155	2	JL0071	alpha-pilin - Moraxel	2.73e+02	186	259	2	C70530	hypothetical protein	2.73e+02
114	157	2	G65156	hypothetical 17.5 kD	2.73e+02	187	260	2	J50341	indole-3-glycerol-pho	2.73e+02
115	157	2	S04915	T-cell receptor delta	2.73e+02	188	260	2	A70040	molybdate-binding pro	2.73e+02
116	158	2	S37355	glutamate--ammonia li	2.73e+02	189	261	2	G69099	probable pyruvate syn	2.73e+02
117	160	2	E34768	ORF3 protein - orf vi	2.73e+02	190	262	2	C70587	hypothetical protein	2.73e+02
118	160	2	H69015	ribosomal protein L10	2.73e+02	191	265	2	C70332	hypothetical protein	2.73e+02
119	161	2	C70613	probable secE - Mycob	2.73e+02	192	265	2	D40362	trpA protein - Methan	2.73e+02
120	164	2	A45512	cold-regulated protei	2.73e+02	193	266	2	JC4850	trypsin-like protein	2.73e+02
121	166	2	F70562	hypothetical protein	2.73e+02	194	267	1	TSPSAA	tryptophan synthase (	2.73e+02
122	170	2	B37074	hypothetical protein	2.73e+02	195	267	2	D70819	probable cyclase invo	2.73e+02
123	174	2	A38487	single-stranded DNA-b	2.73e+02	196	268	1	TSECA	tryptophan synthase (	2.73e+02
124	174	2	A70651	hypothetical protein	2.73e+02	197	268	1	TSEBAT	tryptophan synthase (	2.73e+02
125	174	2	G64367	Wilm's tumor suppress	2.73e+02	198	270	2	D69791	conserved hypothetica	2.73e+02
126	175	2	B69417	ubiquinol-cytochrome	2.73e+02	199	272	2	D70312	hypothetical protein	2.73e+02
127	175	1	DBSCIB	single-stranded DNA-b	2.73e+02	200	273	2	A69089	tryptophan synthase,	2.73e+02
128	179	2	I40109	outer surface protein	2.73e+02	201	274	2	B70684	hypothetical protein	2.73e+02
129	179	2	S54222	outer surface protein	2.73e+02	202	274	2	S35339	hypothetical protein	2.73e+02
130	183	2	S61839	hypothetical protein	2.73e+02	203	275	2	D70737	hypothetical protein	2.73e+02
131	185	2	B70072	conserved hypothetica	2.73e+02	204	278	2	D64619	outer membrane protei	2.73e+02
132	187	2	S39894	probable iron-sulfur	2.73e+02	205	279	2	JH0655	macrotetrolide resist	2.73e+02
133	188	2	T02822	probable membrane pro	2.73e+02	206	282	2	S44825	F54F2.8 protein - Cae	2.73e+02
134	191	2	A71985	hypothetical protein	2.73e+02	207	283	2	B71086	probable 3-methyl-2-o	2.73e+02
135	194	2	A70609	hypothetical protein	2.73e+02	208	283	2	A69207	dihydrodipicolinate s	2.73e+02
136	194	2	S70277	outer surface protein	2.73e+02	209	285	2	D70852	hypothetical protein	2.73e+02
137	197	2	S39529	probable p23 protein	2.73e+02	210	285	2	B71653	hypothetical protein	2.73e+02
138	198	2	S73857	arginine delaminase (E	2.73e+02	211	286	2	G65082	hypothetical protein	2.73e+02
139	198	2	C33465	lic-1 protein C - Hae	2.73e+02	212	286	2	D71502	probable myristoyl gl	2.73e+02
140	199	2	C69659	molybdopterin-guanine	2.73e+02	213	287	2	G71419	hypothetical protein	2.73e+02
141	201	2	JQ0949	proteinase inhibitor	2.73e+02	214	288	2	B24372	DnII endonuclease (d	2.73e+02
142	202	1	OPRUE	glutathione peroxidase	2.73e+02	215	292	2	S03421	T-cell receptor delta	2.73e+02
143	205	2	S54040	hypothetical protein	2.73e+02	216	293	2	C71267	probable carboxyleste	2.73e+02
144	206	2	A29581	alpha-amylase/trypsin	2.73e+02	217	293	2	A40131	T-cell receptor delta	2.73e+02
145	206	2	JS0646	22K antifungal protei	2.73e+02	218	294	2	A32377	lymphocyte surface an	2.73e+02
146	211	2	T03685	probable superoxide d	2.73e+02	219	295	2	B32058	mcbB protein - Escher	2.73e+02
147	211	2	S69928	outer surface protein	2.73e+02	220	298	2	I64138	adhesin homolog HI173	2.73e+02
148	213	2	A34091	phosphoribosylanthran	2.73e+02	221	298	2	A46301	ul snRNP-specific pro	2.73e+02
149	216	2	S09509	outer membrane protei	2.73e+02	222	299	2	B55527	pqqG protein - Methyl	2.73e+02
150	219	2	I52644	neuronal protein - ra	2.73e+02	223	300	2	A42204	phosphoenolpyruvate m	2.73e+02
151	219	2	A69505	uridylylate kinase (pyr	2.73e+02	224	301	2	G69961	phosphoenolpyruvate m	2.73e+02
152	219	2	S47083	dnaK-type molecular c	2.73e+02	225	302	2	S52142	glycosyl transferase a	2.73e+02
153	220	2	S44966	lmbR protein - Strept	2.73e+02	226	301	2	H65113	hypothetical 30.8 kD	2.73e+02
154	222	2	S70878	galactosyl transferase	2.73e+02	227	305	1	R5D0P0	ribosomal protein p0	2.73e+02
155	222	2	S76133	sensory transduction	2.73e+02	228	305	2	T03558	ribose transport ATP-	2.73e+02
156	223	2	I59173	glutamate decarboxyla	2.73e+02	229	309	2	S72881	dihydroorotate oxidas	2.73e+02
157	227	2	T02075	antifungal zeamatin-1	2.73e+02	230	309	1	JNECFS	terminator sequence-b	2.73e+02
158	228	2	F69038	orotidine 5' monophos	2.73e+02	231	310	2	JL0091	beta-lactamase (EC 3.	2.73e+02
159	231	2	D45934	phosphoglycolate phos	2.73e+02	232	311	2	S02714	beta-lactamase (EC 3.	2.73e+02
160	236	1	WZBEK1	IR5 protein - equine	2.73e+02	233	313	2	G69498	calcium-binding prote	2.73e+02
161	236	2	F69314	conserved hypothetica	2.73e+02	234	313	2	E64981	penicillin-binding pr	2.73e+02
162	238	1	WZBEF8	gene 66 protein - equ	2.73e+02	235	313	2	S27698	phosphoenolpyruvate m	2.73e+02
163	238	2	E65113	hypothetical protein	2.73e+02	236	313	2	H71646	hypothetical protein	2.73e+02
164	241	2	D71120	hypothetical protein	2.73e+02	237	313	2	A35132	metalloproteinase (EC	2.73e+02
165	241	2	F70365	pseudouridine synthas	2.73e+02	238	314	2	F70505	probable trna delt-	2.73e+02
166	241	2	F64865	murein hydrolase E (E	2.73e+02	239	314	2	T01554	hypothetical protein	2.73e+02
167	241	2	S65536	hypothetical 28.3K pr	2.73e+02	240	315	2	A26044	prephenate dehydratas	2.73e+02
168	242	2	E59002	conserved hypothetica	2.73e+02	241	315	1	AJY1NB	glutamate--ammonia li	2.73e+02
169	244	2	B47468	cytochrome-c oxidase	2.73e+02	242	316	1	PNBSLC	beta-lactamase (EC 3.	2.73e+02

243	5	16.1	317	2	S41807	glucose-6-phosphate i	2.73e+02	316	5	16.1	356	2	C70683	probable subI - Mycob	2.73e+02
244	5	16.1	318	2	JC6059	1-aminocyclopropane-1	2.73e+02	317	5	16.1	356	2	S52041	Gln 1.2 protein - rad	2.73e+02
245	5	16.1	318	2	C70661	probable moew - Mycob	2.73e+02	318	5	16.1	356	2	JN0041	glutamate--ammonia li	2.73e+02
246	5	16.1	319	2	A70959	probable aspartate ca	2.73e+02	319	5	16.1	357	1	JRZQ8	glutamate--ammonia li	2.73e+02
247	5	16.1	320	2	S22458	ethylene-forming enzy	2.73e+02	320	5	16.1	357	1	HLMS37	MHC class I histocomp	2.73e+02
248	5	16.1	321	2	T02754	probable 1-aminocyclo	2.73e+02	321	5	16.1	357	2	S62712	glutamate--ammonia li	2.73e+02
249	5	16.1	322	2	S52712	1-aminocyclopropane-1	2.73e+02	322	5	16.1	357	2	S39477	glutamate--ammonia li	2.73e+02
250	5	16.1	323	1	PRJLHD	protease (EC 3.4.23	2.73e+02	323	5	16.1	357	2	S12169	glutamate--ammonia li	2.73e+02
251	5	16.1	326	2	S39852	phosphate-binding pro	2.73e+02	324	5	16.1	357	1	AJPMQA	isopenicillin N acylt	2.73e+02
252	5	16.1	326	2	I40790	acetoin dehydrogenase	2.73e+02	325	5	16.1	357	2	S33321	glutamate--ammonia li	2.73e+02
253	5	16.1	328	2	I51215	inhibin alpha-subunit	2.73e+02	326	5	16.1	358	2	A34907	plasma membrane glyco	2.73e+02
254	5	16.1	328	1	ZHRPFA	gene H protein - phag	2.73e+02	327	5	16.1	358	1	AJLCOB	glutamate--ammonia li	2.73e+02
255	5	16.1	328	1	DEHEU7	estradiol 17beta-dehy	2.73e+02	328	5	16.1	359	2	S24240	lymphocyte surface an	2.73e+02
256	5	16.1	328	2	H71871	hypothetical protein	2.73e+02	329	5	16.1	361	2	A42376	lymphocyte surface an	2.73e+02
257	5	16.1	328	2	S52222	DNA-directed RNA poly	2.73e+02	330	5	16.1	361	2	JH0417	cell adhesion molecuol	2.73e+02
258	5	16.1	331	2	JCS442	cathepsin L-like cyst	2.73e+02	331	5	16.1	361	2	G02251	cell surface glycopro	2.73e+02
259	5	16.1	331	2	JCS441	cathepsin L-like cyst	2.73e+02	332	5	16.1	361	2	A45211	prostaglandin E recep	2.73e+02
260	5	16.1	331	2	G69830	lipotein-protein ligas	2.73e+02	333	5	16.1	361	2	A48373	high-alkaline serine	2.73e+02
261	5	16.1	332	1	QRCDI1	iron(III) dicitrate t	2.73e+02	334	5	16.1	362	2	A30901	lymphocyte adhesion r	2.73e+02
262	5	16.1	332	2	JQ3367	hypothetical protein	2.73e+02	335	5	16.1	362	2	A35616	T-cell surface glycop	2.73e+02
263	5	16.1	332	2	S60935	hypothetical protein	2.73e+02	336	5	16.1	362	2	S48689	prostaglandin E(2) re	2.73e+02
264	5	16.1	332	2	S11714	transcription initiat	2.73e+02	337	5	16.1	362	2	S22395	fecuin precursor - pi	2.73e+02
265	5	16.1	333	2	F64406	N5-methyl-tetrahydrom	2.73e+02	338	5	16.1	363	2	A37009	CD44 homolog membrane	2.73e+02
266	5	16.1	333	2	S74876	phosphate-binding per	2.73e+02	339	5	16.1	364	1	PAQXF	fructose-bisphosphata	2.73e+02
267	5	16.1	334	2	JE0231	metalloendopeptidase	2.73e+02	340	5	16.1	364	2	S65009	prostaglandin E recep	2.73e+02
268	5	16.1	336	2	D64589	quinolinate synthetas	2.73e+02	341	5	16.1	364	2	JC2115	prostaglandin E recep	2.73e+02
269	5	16.1	336	2	A71828	quinolinate synthetas	2.73e+02	342	5	16.1	365	2	A71005	probable cell divisio	2.73e+02
270	5	16.1	337	2	C70473	phosphate-binding per	2.73e+02	343	5	16.1	365	2	A34424	CD44 membrane glycop	2.73e+02
271	5	16.1	338	2	C42452	acetoin-2,6-dichlorop	2.73e+02	344	5	16.1	365	2	I72170	MHC class I histocomp	2.73e+02
272	5	16.1	338	2	JCS443	cathepsin L-like cyst	2.73e+02	345	5	16.1	365	2	I72171	HLA-Aw33.1, HLA-Aw43	2.73e+02
273	5	16.1	339	2	B71951	aliphatic amidase - H	2.73e+02	346	5	16.1	365	2	A45847	MHC class I histocomp	2.73e+02
274	5	16.1	339	2	F64556	aliphatic amidase - H	2.73e+02	347	5	16.1	365	2	A42414	prostaglandin E recep	2.73e+02
275	5	16.1	340	1	AJSM2H	glutamate--ammonia li	2.73e+02	348	5	16.1	365	1	AJFFC2	glutamate--ammonia li	2.73e+02
276	5	16.1	341	2	I55623	thromboxane A2 recept	2.73e+02	349	5	16.1	365	2	JN0693	prostaglandin E2 rece	2.73e+02
277	5	16.1	341	2	JH0606	thromboxane A2 recept	2.73e+02	350	5	16.1	365	2	I38610	MHC class I histocomp	2.73e+02
278	5	16.1	341	2	S26206	hypothetical protein	2.73e+02	351	5	16.1	365	2	JC4027	glutamate--ammonia li	2.73e+02
279	5	16.1	342	2	S62711	glutamate--ammonia li	2.73e+02	352	5	16.1	366	2	A53286	cell-surface glycopro	2.73e+02
280	5	16.1	343	1	AJSM2V	glutamate--ammonia li	2.73e+02	353	5	16.1	366	2	S51280	EP3-alpha receptor -	2.73e+02
281	5	16.1	344	2	S22796	probable portal prote	2.73e+02	354	5	16.1	367	2	C69910	N-acetylmuramoyl-L-al	2.73e+02
282	5	16.1	345	2	JC1174	amidase (EC 3.5.1.4),	2.73e+02	355	5	16.1	367	2	F71160	hypothetical protein	2.73e+02
283	5	16.1	346	2	A26741	amidase (EC 3.5.1.4),	2.73e+02	356	5	16.1	368	2	S39478	glutamate--ammonia li	2.73e+02
284	5	16.1	347	3	T02669	hypothetical protein	2.73e+02	357	5	16.1	370	2	A69396	ribose ABC transporte	2.73e+02
285	5	16.1	347	2	S60652	synapse-associated pr	2.73e+02	358	5	16.1	371	2	D71063	hypothetical protein	2.73e+02
286	5	16.1	348	2	I37271	cyllicin II - human	2.73e+02	359	5	16.1	371	1	WMBEY0	dUTP pyrophosphatase	2.73e+02
287	5	16.1	350	2	C70072	serine/threonine prot	2.73e+02	360	5	16.1	371	2	S49786	3-isopropylmalate deh	2.73e+02
288	5	16.1	351	2	S45305	CD44 antigen precurs	2.73e+02	361	5	16.1	372	2	A64462	8-amino-7-oxononanoat	2.73e+02
289	5	16.1	351	2	S43987	replication-associate	2.73e+02	362	5	16.1	372	2	C39364	GDP-1 embryonic growt	2.73e+02
290	5	16.1	351	2	H70682	probable cysA - Mycob	2.73e+02	363	5	16.1	372	2	I40575	histidine protein kin	2.73e+02
291	5	16.1	351	2	S60653	synapse associated pr	2.73e+02	364	5	16.1	373	2	A69158	sensory transduction	2.73e+02
292	5	16.1	351	2	JU0056	flagellin - Serratia	2.73e+02	365	5	16.1	373	2	S39722	spore coat polysaccha	2.73e+02
293	5	16.1	351	2	B24356	replication-associate	2.73e+02	366	5	16.1	374	2	S18602	glutamate--ammonia li	2.73e+02
294	5	16.1	352	2	F64529	translocation releasing	2.73e+02	367	5	16.1	374	2	S18601	glutamate--ammonia li	2.73e+02
295	5	16.1	352	2	F71977	peptide chain release	2.73e+02	368	5	16.1	374	2	S18603	glutamate--ammonia li	2.73e+02
296	5	16.1	352	1	AJRXQ	glutamate--ammonia li	2.73e+02	369	5	16.1	375	2	S03569	glutamate--ammonia li	2.73e+02
297	5	16.1	353	2	E64023	hypothetical protein	2.73e+02	370	5	16.1	375	2	T03593	probable anthocyanidi	2.73e+02
298	5	16.1	354	1	S52040	Gln 1.1 protein - rad	2.73e+02	371	5	16.1	377	1	PWNTG	H+-transporting ATP s	2.73e+02
299	5	16.1	354	1	VVVP1J	coat protein VP1 - po	2.73e+02	372	5	16.1	378	2	D69658	acyl-CoA dehydrogenas	2.73e+02
300	5	16.1	355	2	S39480	glutamate--ammonia li	2.73e+02	373	5	16.1	379	2	B70579	probable cell divisio	2.73e+02
301	5	16.1	355	2	JQ0937	glutamate--ammonia li	2.73e+02	374	5	16.1	380	2	E70933	hypothetical protein	2.73e+02
302	5	16.1	355	1	AJPMQ1	glutamate--ammonia li	2.73e+02	375	5	16.1	380	2	JC5747	beta-ketoacylsynthase	2.73e+02
303	5	16.1	356	2	S40110	glutamate--ammonia li	2.73e+02	376	5	16.1	381	2	A40481	RNA-directed RNA poly	2.73e+02
304	5	16.1	356	2	T03255	glutamate--ammonia li	2.73e+02	377	5	16.1	381	2	B71926	cag island protein -	2.73e+02
305	5	16.1	356	2	T03253	glutamate--ammonia li	2.73e+02	378	5	16.1	381	2	G64047	cystathionine gamma-1	2.73e+02
306	5	16.1	356	2	S71580	glutamate--ammonia li	2.73e+02	379	5	16.1	382	2	I39695	cephalosporinase (EC	2.73e+02
307	5	16.1	356	1	AJRZQG	glutamate--ammonia li	2.73e+02	380	5	16.1	382	2	A70078	conserved hypothetica	2.73e+02
308	5	16.1	356	1	AJFQBQ	glutamate--ammonia li	2.73e+02	381	5	16.1	383	2	G44667	NA+/H+ antiporter - H	2.73e+02
309	5	16.1	356	2	S71579	glutamate--ammonia li	2.73e+02	382	5	16.1	383	2	H71848	probable na+/h+ antip	2.73e+02
310	5	16.1	356	1	AJFBO	glutamate--ammonia li	2.73e+02	383	5	16.1	384	2	G70961	hypothetical protein	2.73e+02
311	5	16.1	356	1	AJRAQ	glutamate--ammonia li	2.73e+02	384	5	16.1	386	2	A41950	retrovirus-related hy	2.73e+02
312	5	16.1	356	2	S52042	Gln 1.3 protein - rad	2.73e+02	385	5	16.1	387	2	E65010	multidrug resistance	2.73e+02
313	5	16.1	356	2	S39479	glutamate--ammonia li	2.73e+02	386	5	16.1	388	2	E70894	probable metB protein	2.73e+02
314	5	16.1	356	2	S49976	glutamate--ammonia li	2.73e+02	387	5	16.1	390	2	A57616	inward rectifier K ch	2.73e+02
315	5	16.1	356	1	AJFBOA	glutamate--ammonia li	2.73e+02	388	5	16.1	390	2	A64549	conserved hypothetica	2.73e+02

389	5	16.1	390	2	J04689	inwardly rectifying p	2.73e-02	462	5	16.1	488	2	A40367	catalase (EC 1.11.1.6	2.73e+02
390	5	16.1	390	2	S68403	inward rectifier pota	2.73e-02	463	5	16.1	489	1	VBETA	glycoprotein A - turk	2.73e+02
391	5	16.1	390	2	T00668	hypothetical protein	2.73e-02	464	5	16.1	490	2	H70473	IMP dehydrogenase (EC	2.73e+02
392	5	16.1	394	2	T00668	polygalacturonase hom	2.73e-02	465	5	16.1	493	2	S13530	CD44E protein, epithe	2.73e+02
393	5	16.1	395	2	I77371	CD44R5 - human	2.73e-02	466	5	16.1	493	2	JC4372	IMP dehydrogenase (EC	2.73e+02
394	5	16.1	395	2	C69292	alcohol dehydrogenase	2.73e-02	467	5	16.1	496	2	S33597	protein kinase chk1 (	2.73e+02
395	5	16.1	398	2	E71379	probable tpr protein	2.73e-02	468	5	16.1	496	2	G64501	IMP dehydrogenase (EC	2.73e+02
396	5	16.1	399	1	AJFFLM	glutamate--ammonia li	2.73e-02	469	5	16.1	497	1	S20174	protein kinase MEK1 (	2.73e+02
397	5	16.1	399	2	B41608	hypothetical protein	2.73e-02	470	5	16.1	497	2	S69545	apoptosis inhibitor I	2.73e+02
398	5	16.1	400	2	A48050	coagulation factor B	2.73e-02	471	5	16.1	498	2	D64878	probable glutamate--a	2.73e+02
399	5	16.1	402	2	S74332	hypothetical protein	2.73e-02	472	5	16.1	499	2	E69123	glutamate synthase (N	2.73e+02
400	5	16.1	404	2	A69049	molybdenum cofactor b	2.73e-02	473	5	16.1	503	2	E38745	cell adhesion molecucl	2.73e+02
401	5	16.1	406	2	F70632	probable metz protein	2.73e-02	474	5	16.1	504	2	A38221	chitinase (EC 3.2.1.1	2.73e+02
402	5	16.1	409	1	A70527	transcription factor	2.73e-02	475	5	16.1	506	2	S76981	probable phosphoester	2.73e+02
403	5	16.1	410	2	C70001	malate dehydrogenase	2.73e-02	476	5	16.1	507	2	S48458	probable membrane pro	2.73e+02
404	5	16.1	413	2	I41060	3-oxoacyl-[acyl-carri	2.73e-02	477	5	16.1	509	2	T02260	beta-fructofuranosida	2.73e+02
405	5	16.1	418	2	E47056	nickel-cobalt resista	2.73e-02	478	5	16.1	509	2	A57512	beta-glucosidase BGQ6	2.73e+02
406	5	16.1	420	2	C70404	diaminopimelate decar	2.73e-02	479	5	16.1	510	1	FWQFA	H+-transporting ATP s	2.73e+02
407	5	16.1	423	2	I38056	nicotinic acetylcholi	2.73e-02	480	5	16.1	512	1	ALBSL	alpha-amylase (EC 3.2	2.73e+02
408	5	16.1	425	2	E69551	thiamine biosynthesis	2.73e-02	481	5	16.1	514	2	B71252	probable tpr protein	2.73e+02
409	5	16.1	426	2	JH0518	lymphocyte homing rec	2.73e-02	482	5	16.1	515	2	S72987	acetylactate synthase	2.73e+02
410	5	16.1	426	2	F64419	hypothetical protein	2.73e-02	483	5	16.1	518	1	A27705	alpha-amylase (EC 3.2	2.73e+02
411	5	16.1	426	2	B24248	inhibin beta-A chain	2.73e-02	484	5	16.1	520	2	A35883	laccase (EC 1.10.3.2)	2.73e+02
412	5	16.1	427	2	I31580	XfRH2 protein - Afric	2.73e-02	485	5	16.1	520	2	B35883	ligninolytic phenolox	2.73e+02
413	5	16.1	428	2	A69085	3-isopropylmalate deh	2.73e-02	486	5	16.1	523	1	A44195	UL21 protein homolog	2.73e+02
414	5	16.1	429	2	S27924	gene lF2 protein - hu	2.73e-02	487	5	16.1	523	1	A60408	glycoprotein gp57-65	2.73e+02
415	5	16.1	430	2	S74749	ABC-type transport pr	2.73e-02	488	5	16.1	529	2	H70736	probable guaB2 protei	2.73e+02
416	5	16.1	431	2	B37802	crtx protein - Erwini	2.73e-02	489	5	16.1	529	2	S72823	IMP dehydrogenase (EC	2.73e+02
417	5	16.1	433	1	J70875	pyrimidine-nucleoside	2.73e-02	490	5	16.1	530	2	S46589	UTR1 protein - yeast	2.73e+02
418	5	16.1	433	2	B69495	aconitase (acon) homol	2.73e-02	491	5	16.1	530	2	C47113	glucuronosyltransfera	2.73e+02
419	5	16.1	434	1	S49457	pyrimidine-nucleoside	2.73e-02	492	5	16.1	535	2	A44475	N-acetylglucosamine	2.73e+02
420	5	16.1	437	2	I39578	nccC protein - Alkali	2.73e-02	493	5	16.1	536	2	A36395	spore wall maturatio	2.73e+02
421	5	16.1	442	2	A57041	transcription regulat	2.73e-02	494	5	16.1	538	2	S36424	cytochrome-c oxidase	2.73e+02
422	5	16.1	445	1	NUBSS	glucose-6-phosphate i	2.73e-02	495	5	16.1	545	2	S51563	heat shock protein gr	2.73e+02
423	5	16.1	445	2	S35937	glucose-6-phosphate i	2.73e-02	496	5	16.1	547	2	B64701	probable 60K inner me	2.73e+02
424	5	16.1	446	2	A38244	citrate transport sys	2.73e-02	497	5	16.1	548	2	JQ1142	Tlrc protein - Strept	2.73e+02
425	5	16.1	446	2	B42661	citrate carrier prote	2.73e-02	498	5	16.1	549	2	E71818	probable inner membra	2.73e+02
426	5	16.1	446	2	A42661	citrate carrier prote	2.73e-02	499	5	16.1	551	2	S66740	probable transcrip	2.73e+02
427	5	16.1	448	2	P50047	serine proteinase (EC	2.73e-02	500	5	16.1	553	2	S77623	mannuronan C-5-epimer	2.73e+02
428	5	16.1	448	2	D71477	hypothetical protein	2.73e-02	501	5	16.1	553	2	H70715	probable glucose-6-ph	2.73e+02
429	5	16.1	449	1	NUBSSA	glucose-6-phosphate i	2.73e-02	502	5	16.1	554	2	H70011	exo-alpha-1,4-glucosi	2.73e+02
430	5	16.1	450	2	T02368	hypothetical protein	2.73e-02	503	5	16.1	554	1	NUBY	glucose-6-phosphate i	2.73e+02
431	5	16.1	450	2	G70310	conserved hypotetica	2.73e-02	504	5	16.1	557	1	TLBPT3	tail fiber protein -	2.73e+02
432	5	16.1	450	1	JN0812	benzene 1,2-dioxygena	2.73e-02	505	5	16.1	557	1	NUVKL	glucose-6-phosphate i	2.73e+02
433	5	16.1	451	2	B59675	glucose-6-phosphate i	2.73e-02	506	5	16.1	560	2	S41808	glucose-6-phosphate i	2.73e+02
434	5	16.1	454	2	S75741	hypothetical protein	2.73e-02	507	5	16.1	566	2	S07508	DNA primase - phase T	2.73e+02
435	5	16.1	455	2	E71074	hypothetical protein	2.73e-02	508	5	16.1	566	2	IDBFA7	DNA primase - phase T	2.73e+02
436	5	16.1	457	2	S34746	anthranilate synthase	2.73e-02	509	5	16.1	567	2	T02094	glucose-6-phosphate i	2.73e+02
437	5	16.1	458	2	T02571	probable myrosinase-b	2.73e-02	510	5	16.1	568	2	S57830	glucose-6-phosphate i	2.73e+02
438	5	16.1	458	2	B39533	transcription factor	2.73e-02	511	5	16.1	569	2	S41806	glucose-6-phosphate i	2.73e+02
439	5	16.1	459	1	B34258	transcription factor	2.73e-02	512	5	16.1	569	2	S57831	glucose-6-phosphate i	2.73e+02
440	5	16.1	459	2	T02569	probable myrosinase-b	2.73e-02	513	5	16.1	569	2	S23542	glucose-6-phosphate i	2.73e+02
441	5	16.1	463	2	S74511	malic enzyme - Synecch	2.73e-02	514	5	16.1	570	1	HNNZAV	hemagglutinin-neurami	2.73e+02
442	5	16.1	465	2	G71931	probable outer membra	2.73e-02	515	5	16.1	571	2	H46328	hemagglutinin-neurami	2.73e+02
443	5	16.1	466	2	A36674	transcription factor	2.73e-02	516	5	16.1	573	2	A43533	cell surface glycopro	2.73e+02
444	5	16.1	468	1	A54258	transcription factor	2.73e-02	517	5	16.1	575	1	VCLJHD	env polyprotein precu	2.73e+02
445	5	16.1	471	2	T02186	hypothetical protein	2.73e-02	518	5	16.1	577	1	D46328	hemagglutinin-neurami	2.73e+02
446	5	16.1	473	2	S07555	hypothetical protein	2.73e-02	519	5	16.1	577	1	E46328	hemagglutinin-neurami	2.73e+02
447	5	16.1	473	2	S40507	endoglucanase - rumen	2.73e-02	520	5	16.1	577	1	HNNZNC	hemagglutinin-neurami	2.73e+02
448	5	16.1	473	2	T03357	forkhead transcrip	2.73e-02	521	5	16.1	585	2	A39286	parathyroid hormone /	2.73e+02
449	5	16.1	474	2	I38240	SOX-4 protein - human	2.73e-02	522	5	16.1	585	2	S61534	glutamate decarboxyla	2.73e+02
450	5	16.1	474	2	E24723	trpc protein - Coryne	2.73e-02	523	5	16.1	591	1	WMBPQ2	gene p2 protein - pha	2.73e+02
451	5	16.1	476	2	A41463	alkaline metalloprote	2.73e-02	524	5	16.1	591	2	S73708	MG321 homolog H08_orf	2.73e+02
452	5	16.1	479	2	A38578	protein kinase 2 (EC	2.73e-02	525	5	16.1	593	2	S51776	glutamate decarboxyla	2.73e+02
453	5	16.1	479	2	A38307	metalloproteinase C (	2.73e-02	526	5	16.1	593	1	A41367	glutamate decarboxyla	2.73e+02
454	5	16.1	479	2	S66999	alkaline metalloprote	2.73e-02	527	5	16.1	593	2	S48135	glutamate decarboxyla	2.73e+02
455	5	16.1	480	2	G64580	hypothetical protein	2.73e-02	528	5	16.1	594	2	S51775	glutamate decarboxyla	2.73e+02
456	5	16.1	481	2	E64623	IMP dehydrogenase (EC	2.73e-02	529	5	16.1	594	2	JC4065	glutamate decarboxyla	2.73e+02
457	5	16.1	481	2	H71890	inosine-5'-monophosph	2.73e-02	530	5	16.1	594	1	A46758	glutamate decarboxyla	2.73e+02
458	5	16.1	482	2	JF0302	cellulase (EC 3.2.1.4	2.73e-02	531	5	16.1	594	1	B41935	glutamate decarboxyla	2.73e+02
459	5	16.1	485	2	JC4998	IMP dehydrogenase (EC	2.73e-02	532	5	16.1	595	2	JQ1684	anthranilate synthase	2.73e+02
460	5	16.1	486	2	E71456	probable inosine-5'-m	2.73e-02	533	5	16.1	596	1	TVFVSA	env-sea polyprotein -	2.73e+02
461	5	16.1	488	2	S32326	IMP dehydrogenase (EC	2.73e-02	534	5	16.1	598	2	D71365	probable tpr protein	2.73e+02

535	5	16.1	598	2	H71362	probable tpr protein	2.73e+02	764	1	BBHU	complement factor B p	2.73e+02
536	5	16.1	599	2	H70143	hypothetical protein	2.73e+02	767	2	F71694	malic enzyme (tme) RP	2.73e+02
537	5	16.1	601	2	A64222	heat shock protein dn	2.73e+02	771	2	A40966	membrane protein SKM1	2.73e+02
538	5	16.1	603	2	B70886	probable penicillin b	2.73e+02	774	2	S13670	basic polymerase 2 -	2.73e+02
539	5	16.1	607	1	NUUTB	glucose-6-phosphate i	2.73e+02	774	1	S3IV50	RNA-directed RNA poly	2.73e+02
540	5	16.1	609	2	H71301	probable tpr protein	2.73e+02	775	2	P65769	maltooligosyl trehalo	2.73e+02
541	5	16.1	612	2	S29318	lysophospholipase (EC	2.73e+02	777	2	A35966	chemotaxis protein fr	2.73e+02
542	5	16.1	615	2	S54468	AlpI protein - yeast	2.73e+02	779	2	H71301	probable membrane-bou	2.73e+02
543	5	16.1	616	1	HNZQD	hemagglutinin-neurami	2.73e+02	780	1	KIHUFM	6-phosphofructokinase	2.73e+02
544	5	16.1	616	1	HNZU1	hemagglutinin-neurami	2.73e+02	780	1	KIRBF	6-phosphofructokinase	2.73e+02
545	5	16.1	616	1	A46328	hemagglutinin-neurami	2.73e+02	780	1	JC2055	6-phosphofructokinase	2.73e+02
546	5	16.1	616	1	T01990	anthranilate synthase	2.73e+02	784	2	JC2055	6-phosphofructokinase	2.73e+02
547	5	16.1	616	1	B46328	hemagglutinin-neurami	2.73e+02	787	2	C36857	B21R protein - variol	2.73e+02
548	5	16.1	616	1	C46328	hemagglutinin-neurami	2.73e+02	791	2	A53206	6-phosphofructokinase	2.73e+02
549	5	16.1	620	2	S30780	hexose metabolism-rel	2.73e+02	792	1	EAHU	elastin precursor, lo	2.73e+02
550	5	16.1	621	2	T05225	hypothetical protein	2.73e+02	801	2	A33630	zinc finger protein, z	2.73e+02
551	5	16.1	621	2	JQ1685	anthranilate synthase	2.73e+02	802	1	A46343	ADES multifunctional	2.73e+02
552	5	16.1	621	2	S27752	anthranilate synthase	2.73e+02	802	2	S48529	NAB3 protein - yeast	2.73e+02
553	5	16.1	621	2	D70762	probable transferase	2.73e+02	804	2	D69650	leucine--tRNA ligase	2.73e+02
554	5	16.1	627	2	E70122	flagellar hook-associ	2.73e+02	807	2	D69102	collagenase - Mechano	2.73e+02
555	5	16.1	628	2	D70722	hypothetical protein	2.73e+02	826	2	S41569	ferric pseudobactin M	2.73e+02
556	5	16.1	631	2	E69115	phage infection prote	2.73e+02	827	1	S10639	fructose phosphotrans	2.73e+02
557	5	16.1	632	2	T00108	hypothetical protein	2.73e+02	827	1	S66094	stage II sporulation	2.73e+02
558	5	16.1	633	2	S49611	probable serine/threo	2.73e+02	829	2	B27211	viRA protein - Agroba	2.73e+02
559	5	16.1	635	2	B64112	dnak-type molecular c	2.73e+02	829	2	S04035	viRA protein - Agroba	2.73e+02
560	5	16.1	635	2	A36868	copA homolog - Xantho	2.73e+02	829	2	S12858	viRA protein - Agroba	2.73e+02
561	5	16.1	638	2	B71904	probable outer membra	2.73e+02	833	2	S01872	viRA protein - Agroba	2.73e+02
562	5	16.1	638	1	IQEDK	dnak-type molecular c	2.73e+02	833	2	S11925	viRA protein - Agroba	2.73e+02
563	5	16.1	638	2	A56695	notch homolog hN prec	2.73e+02	836	2	A69550	hypothetical protein	2.73e+02
564	5	16.1	640	1	HMIVEC	hemagglutinin precurs	2.73e+02	836	2	D64988	yejO protein - Escher	2.73e+02
565	5	16.1	640	1	HMIVEA	hemagglutinin precurs	2.73e+02	837	2	E70835	hypothetical glycine-	2.73e+02
566	5	16.1	640	1	HMIVEB	hemagglutinin precurs	2.73e+02	838	2	S73687	valine--tRNA ligase (	2.73e+02
567	5	16.1	643	2	T03518	hypothetical protein	2.73e+02	857	2	S01864	lipoygenase (EC 1.13	2.73e+02
568	5	16.1	643	2	S59412	sefA protein - Escher	2.73e+02	861	2	S34730	1,4-alpha-glucan bran	2.73e+02
569	5	16.1	649	2	S53126	dnak-type molecular c	2.73e+02	877	2	S76394	hypothetical protein	2.73e+02
570	5	16.1	649	1	TVVPMK	large T antigen - mou	2.73e+02	881	1	W2BE55	gene 55 protein - hum	2.73e+02
571	5	16.1	654	1	HMIVC8	hemagglutinin precurs	2.73e+02	891	2	JN0867	peroxinectin-like pro	2.73e+02
572	5	16.1	655	2	S07412	hemagglutinin-esteras	2.73e+02	894	2	PN0667	peroxinectin-like pro	2.73e+02
573	5	16.1	657	2	T07079	probable ctsp protein	2.73e+02	894	1	S5B1MX	leucine--tRNA ligase	2.73e+02
574	5	16.1	658	2	A44272	photomorphogenesis re	2.73e+02	895	2	A55514	pyruvate dehydrogenas	2.73e+02
575	5	16.1	658	2	JC4598	phosphotransferase sy	2.73e+02	899	2	G02428	prohormone convertase	2.73e+02
576	5	16.1	661	2	S44773	C2984.5 protein - Cae	2.73e+02	905	2	C70758	probable cation trans	2.73e+02
577	5	16.1	663	2	B70460	excinuclease ABC (EC	2.73e+02	907	2	S54353	inter-alpha-inhibitor	2.73e+02
578	5	16.1	664	2	S61231	cycK protein - Rhizob	2.73e+02	911	2	S46497	ak-hsdh bifunctional	2.73e+02
579	5	16.1	667	2	C70719	probable carbamoyl-ph	2.73e+02	912	2	JC6148	subtilisin-like propr	2.73e+02
580	5	16.1	667	2	T07082	hypothetical glycine-	2.73e+02	915	2	A48225	subtilisin-like propr	2.73e+02
581	5	16.1	667	2	S11840	viD3 protein - Agrob	2.73e+02	915	2	B48225	probable proporotein c	2.73e+02
582	5	16.1	673	2	B70895	hypothetical protein	2.73e+02	932	2	JC5953	inter-alpha-inhibitor	2.73e+02
583	5	16.1	675	2	T01112	photomorphogenesis re	2.73e+02	934	2	H71274	probable ankyrin - sy	2.73e+02
584	5	16.1	676	2	S69783	outer membrane protei	2.73e+02	949	2	T02083	H+-transporting ATPas	2.73e+02
585	5	16.1	678	2	S29123	NADPH--ferrihemoprote	2.73e+02	956	2	S50751	H+-transporting ATPas	2.73e+02
586	5	16.1	680	2	S29682	DNA topoisomerase (Ar	2.73e+02	956	2	A45506	H+-transporting ATPas	2.73e+02
587	5	16.1	693	2	K71302	probable tpr protein	2.73e+02	956	2	A43637	H+-transporting ATPas	2.73e+02
588	5	16.1	698	1	IKES5B	colicin V secretion p	2.73e+02	957	2	A41779	H+-transporting ATPas	2.73e+02
589	5	16.1	699	2	I37369	epican - human	2.73e+02	968	2	S46992	protein p130 - rat	2.73e+02
590	5	16.1	704	2	S46000	probable membrane pro	2.73e+02	994	1	SYNCLM	leucine--tRNA ligase	2.73e+02
591	5	16.1	709	2	S38241	hypothetical protein	2.73e+02	996	2	S42208	NAD+ ADP-ribosyltrans	2.73e+02
592	5	16.1	711	2	C40046	antibiotic transport-	2.73e+02	997	2	I39739	mannuronan C5 epimer	2.73e+02
593	5	16.1	713	2	JC2522	nuclear autoantigen -	2.73e+02	997	2	S77625	mannuronan C-5-epimer	2.73e+02
594	5	16.1	714	2	S56893	hypothetical protein	2.73e+02	999	1	IJHUG3	desmoglein 3 precurs	2.73e+02
595	5	16.1	714	2	T09807	hypothetical glycine-	2.73e+02	1008	2	S73711	probable lipoprotein	2.73e+02
596	5	16.1	714	2	S76082	hypothetical protein	2.73e+02	1008	2	F71727	acriflavin resistance	2.73e+02
597	5	16.1	720	2	S61143	KRE6 protein - yeast	2.73e+02	1009	2	S26840	retrovirus-related po	2.73e+02
598	5	16.1	729	2	C64854	ferric-coprogen recep	2.73e+02	1009	2	S64734	retrovirus-related po	2.73e+02
599	5	16.1	731	2	A55229	ornithine decarboxyla	2.73e+02	1015	2	JC5062	phogrin precursor - h	2.73e+02
600	5	16.1	736	1	M4XR3D	structural protein u2	2.73e+02	1015	2	JC5263	transmembrane tyrosin	2.73e+02
601	5	16.1	736	2	S23654	Mu2 protein - reoviru	2.73e+02	1017	2	B70985	probable polyketide s	2.73e+02
602	5	16.1	742	2	A47195	lymphocyte homing rec	2.73e+02	1026	2	A48995	paracrystalline surfa	2.73e+02
603	5	16.1	742	2	S40622	methylmalonyl-CoA mut	2.73e+02	1028	2	T03516	probable outer membra	2.73e+02
604	5	16.1	752	1	C2HU	complement C2 precurs	2.73e+02	1036	2	A29832	HPI layer surface pro	2.73e+02
605	5	16.1	752	2	G70751	probable ctsp protein	2.73e+02	1038	1	MWRACB	myosin beta heavy cha	2.73e+02
606	5	16.1	756	2	S43085	DNA mismatch repair p	2.73e+02	1043	1	A29483	myosin heavy chain I,	2.73e+02
607	5	16.1	760	2	D69865	conserved hypotheticala	2.73e+02	1049	1	S51784	toxin III - Actinobac	2.73e+02

681	5	16.1	1052	1	B49219	2.73e-02	toxin III - Actinobac	754	5	16.1	2344	1	RRWRH	genome polyprotein -	2.73e+02
682	5	16.1	1077	2	S70120	2.73e-02	hypothetical protein	755	5	16.1	2344	2	S64740	genome polyprotein -	2.73e+02
683	5	16.1	1094	2	F70697	2.73e-02	probable arabinosyltr	756	5	16.1	2344	2	S55399	genome polyprotein -	2.73e+02
684	5	16.1	1096	2	S62358	2.73e-02	inositol 1,4,5-trisph	757	5	16.1	2395	2	S50820	surface protein type	2.73e+02
685	5	16.1	1100	2	H71616	2.73e-02	SERA antigen/papain-1	758	5	16.1	2399	2	H71879	toxin-like outer memb	2.73e+02
686	5	16.1	1101	2	S15271	2.73e-02	endoglucanase cenC -	759	5	16.1	2425	2	D69426	surface layer protein	2.73e+02
687	5	16.1	1101	2	S51823	2.73e-02	myosin heavy chain AT	760	5	16.1	2471	2	A49128	cell fate determining	2.73e+02
688	5	16.1	1104	2	A36866	2.73e-02	microbial collagenase	761	5	16.1	2514	1	MNWN2	toxin-like outer memb	2.73e+02
689	5	16.1	1113	2	E64215	2.73e-02	hypothetical protein	762	5	16.1	2529	2	B64635	toxin-like outer memb	2.73e+02
690	5	16.1	1127	2	D70671	2.73e-02	pyruvate carboxylase	763	5	16.1	2703	2	A24420	notch protein - fruit	2.73e+02
691	5	16.1	1143	2	S46122	2.73e-02	SNF2 protein homolog	764	5	16.1	3005	2	S33642	homeotic protein zfh-	2.73e+02
692	5	16.1	1174	2	S28976	2.73e-02	DNA-directed RNA poly	765	5	16.1	3016	2	S77300	hypothetical protein -	2.73e+02
693	5	16.1	1176	2	A49848	2.73e-02	nitrite reductase (NA	766	5	16.1	3054	1	GNBEV	genome polyprotein -	2.73e+02
694	5	16.1	1176	2	A27826	2.73e-02	DNA-directed RNA poly	767	5	16.1	3106	1	S33868	laminin alpha-2 chain	2.73e+02
695	5	16.1	1192	2	G70513	2.73e-02	probable meth - Mycob	768	5	16.1	3566	2	A40701	tenascin-X precursor	2.73e+02
696	5	16.1	1197	2	G65010	2.73e-02	sensor protein evgs (	769	5	16.1	3573	2	S23070	erythronolide synthas	2.73e+02
697	5	16.1	1246	2	J00406	2.73e-02	hypothetical protein	770	4	12.9	25	2	S03500	T-cell receptor gamma	4.71e+03
698	5	16.1	1247	2	A33812	2.73e-02	interphotoreceptor re	771	4	12.9	45	1	GENW3B	globin, minor monomer	4.71e+03
699	5	16.1	1264	2	A36858	2.73e-02	G2R protein - variola	772	4	12.9	59	2	I59459	histone H1 - human (f	4.71e+03
700	5	16.1	1286	2	S38058	2.73e-02	hypothetical protein	773	4	12.9	83	2	S65144	pollen allergen group	4.71e+03
701	5	16.1	1287	2	B53739	2.73e-02	vacuolating cytotoxin	774	4	12.9	90	2	S01732	DNA-binding protein H	4.71e+03
702	5	16.1	1287	2	A43488	2.73e-02	genome polyprotein -	775	4	12.9	91	2	S41663	beta-microseminoprote	4.71e+03
703	5	16.1	1288	2	E71884	2.73e-02	vacuolating cytotoxin	776	4	12.9	95	1	R5BS23	ribosomal protein L23	4.71e+03
704	5	16.1	1290	2	G64630	2.73e-02	vacuolating cytotoxin	777	4	12.9	102	2	S04565	histone H2A - chicken	4.71e+03
705	5	16.1	1291	2	A44983	2.73e-02	vacuolating cytotoxin	778	4	12.9	102	2	JC2344	heat shock protein 90	4.71e+03
706	5	16.1	1298	1	EBE75	2.73e-02	immediate-early prote	779	4	12.9	106	1	R5BY1B	acidic ribosomal prot	4.71e+03
707	5	16.1	1300	2	T00317	2.73e-02	serine proteinase esp	780	4	12.9	108	2	E69358	conserved hypothetica	4.71e+03
708	5	16.1	1306	2	A70934	2.73e-02	hypothetical glycine-	781	4	12.9	120	2	S29278	hypothetical protein	4.71e+03
709	5	16.1	1329	2	E70917	2.73e-02	hypothetical glycine-	782	4	12.9	122	2	S16804	histidine triad prote	4.71e+03
710	5	16.1	1346	2	G71613	2.73e-02	hypothetical protein	783	4	12.9	124	2	S11314	histone H2A - polycha	4.71e+03
711	5	16.1	1369	2	S70713	2.73e-02	protein-tyrosine kina	784	4	12.9	134	2	E71945	hypothetical protein	4.71e+03
712	5	16.1	1374	2	S70712	2.73e-02	protein-tyrosine kina	785	4	12.9	136	2	JS0740	H+-transporting ATP s	4.71e+03
713	5	16.1	1403	2	S77624	2.73e-02	mannuronan C-5-epime	786	4	12.9	136	2	F69870	general stress protei	4.71e+03
714	5	16.1	1404	2	A48196	2.73e-02	protein-tyrosine kina	787	4	12.9	137	2	B64938	hypothetical protein	4.71e+03
715	5	16.1	1405	1	DZ2PA	2.73e-02	DNA-directed DNA poly	788	4	12.9	137	2	S35752	dehydrin 7 - barley	4.71e+03
716	5	16.1	1456	2	JQ2294	2.73e-02	hypothetical 16S.IK p	789	4	12.9	138	1	MMBE11	14.6K membrane protei	4.71e+03
717	5	16.1	1464	2	S59856	2.73e-02	collagen alpha 1(III)	790	4	12.9	139	2	S78312	ribulose-bisphosphate	4.71e+03
718	5	16.1	1481	2	S78373	2.73e-02	DNA-directed RNA poly	791	4	12.9	141	2	A24625	hemoglobin alpha-A ch	4.71e+03
719	5	16.1	1487	2	S48719	2.73e-02	phospholipase-A(2) re	792	4	12.9	141	1	HAGRAR	hemoglobin alpha-A ch	4.71e+03
720	5	16.1	1505	2	S28079	2.73e-02	DNA-directed DNA poly	793	4	12.9	155	2	A60130	acidic fibroblast gro	4.71e+03
721	5	16.1	1509	1	A27224	2.73e-02	myosin heavy chain II	794	4	12.9	155	2	I47139	Ig heavy chain variab	4.71e+03
722	5	16.1	1548	2	S4583	2.73e-02	serine proteinase (EC	795	4	12.9	173	2	F45352	orf6 protein - porcin	4.71e+03
723	5	16.1	1584	2	T00026	2.73e-02	brain-specific angiot	796	4	12.9	181	2	E31482	H+-transporting ATP s	4.71e+03
724	5	16.1	1616	2	G70668	2.73e-02	polyketide synthase p	797	4	12.9	182	2	D64017	hypothetical protein	4.71e+03
725	5	16.1	1624	2	C70867	2.73e-02	probable Helix-turn-h	798	4	12.9	188	2	A49020	calcium-binding prote	4.71e+03
726	5	16.1	1643	2	D71630	2.73e-02	outer membrane protei	799	4	12.9	192	2	C64767	yail protein - Escher	4.71e+03
727	5	16.1	1645	2	JN0896	2.73e-02	crystalline surface 1	800	4	12.9	192	2	S22981	T-cell receptor CD3 e	4.71e+03
728	5	16.1	1660	2	A70869	2.73e-02	hypothetical glycine-	801	4	12.9	192	2	JC4663	T-cell receptor CD3 e	4.71e+03
729	5	16.1	1680	2	A43434	2.73e-02	furin (EC 3.4.21.75)	802	4	12.9	193	2	I50676	gene Rem-1 protein -	4.71e+03
730	5	16.1	1723	2	S58880	2.73e-02	receptor DEC-205 - mo	803	4	12.9	195	2	S24296	chorion protein - sil	4.71e+03
731	5	16.1	1751	1	MMHMF	2.73e-02	laminin alpha-2 chain	804	4	12.9	196	2	S47088	psaD protein - Chlamy	4.71e+03
732	5	16.1	1763	1	RRWUF	2.73e-02	genome polyprotein -	805	4	12.9	196	1	A28944	flagellin A1 precursor	4.71e+03
733	5	16.1	1839	2	S77626	2.73e-02	mannuronan C-5-epime	806	4	12.9	205	2	JQ2247	photosystem I chain D	4.71e+03
734	5	16.1	1854	2	I73956	2.73e-02	kinase-related protei	807	4	12.9	207	2	A60695	photosystem I chain I	4.71e+03
735	5	16.1	1901	2	F70806	2.73e-02	hypothetical glycine-	808	4	12.9	208	2	S19605	hydrogenase regulator	4.71e+03
736	5	16.1	1934	1	I48153	2.73e-02	beta-myosin heavy cha	809	4	12.9	208	2	S00449	photosystem I chain I	4.71e+03
737	5	16.1	1935	1	S06006	2.73e-02	myosin beta heavy cha	810	4	12.9	209	2	G64643	hypothetical protein	4.71e+03
738	5	16.1	1935	1	A37102	2.73e-02	myosin beta heavy cha	811	4	12.9	209	2	S01988	late L3 23K proteins	4.71e+03
739	5	16.1	1951	2	B43963	2.73e-02	RNA viral polymerase	812	4	12.9	215	2	I51087	gene HMG-T2 protein -	4.71e+03
740	5	16.1	1956	2	T00051	2.73e-02	hypothetical protein	813	4	12.9	225	2	G70340	Mg(2+) transport ATPa	4.71e+03
741	5	16.1	2024	2	A54103	2.73e-02	centrosome autoantige	814	4	12.9	225	2	S48780	L1 protein - human pa	4.71e+03
742	5	16.1	2101	2	S57245	2.73e-02	insulin receptor homo	815	4	12.9	241	2	S77596	cytochrome-c oxidase	4.71e+03
743	5	16.1	2103	1	J01621	2.73e-02	genome polyprotein -	816	4	12.9	242	2	A39164	homeotic protein Hox	4.71e+03
744	5	16.1	2126	2	E70522	2.73e-02	probable polyketide s	817	4	12.9	245	2	D36145	precorrin-2 methyltra	4.71e+03
745	5	16.1	2148	2	A56081	2.73e-02	insulin receptor - fr	818	4	12.9	248	1	TVWVRS	transforming protein	4.71e+03
746	5	16.1	2150	2	S13553	2.73e-02	hypothetical protein	819	4	12.9	252	2	C69022	methyl coenzyme M red	4.71e+03
747	5	16.1	2151	1	S16449	2.73e-02	genome polyprotein -	820	4	12.9	255	2	A35255	chlorocatechol 1,2-di	4.71e+03
748	5	16.1	2156	1	RRVUNE	2.73e-02	genome polyprotein -	821	4	12.9	258	1	OLECA	heat-labile enterotox	4.71e+03
749	5	16.1	2180	2	A47651	2.73e-02	zinc-finger protein -	822	4	12.9	259	2	A64509	tRNA-pseudouridine sy	4.71e+03
750	5	16.1	2210	1	RRXPTV	2.73e-02	genome polyprotein -	823	4	12.9	261	1	JC1293	serine O-acetyltransf	4.71e+03
751	5	16.1	2291	2	S11238	2.73e-02	polymerase - Berne vi	824	4	12.9	261	2	S35720	genome polyprotein -	4.71e+03
752	5	16.1	2317	2	I56752	2.73e-02	c-ros-1 tyrosine kina	825	4	12.9	273	2	I40102	outer surface protein	4.71e+03
753	5	16.1	2338	2	I73957	2.73e-02	c-ros-1 tyrosine kina	826	4	12.9	274	2	F70642	probable rp5c protein	4.71e+03



827	4	12.9	274	2	I51650	gene Xcad-3 protein -	4.71e+03	900	4	12.9	451	2	C49898	cellobiose phosphotra	4.71e+03
828	4	12.9	276	2	G64333	hypothetical protein	4.71e+03	901	4	12.9	452	2	S47236	ribulose-bisphosphate	4.71e+03
829	4	12.9	279	2	JC4327	uridine phosphorylase	4.71e+03	902	4	12.9	459	2	H64667	conserved hypotheticala	4.71e+03
830	4	12.9	285	2	S68332	dihydropteroate synth	4.71e+03	903	4	12.9	461	2	E64630	virulence factor mv1N	4.71e+03
831	4	12.9	288	2	G68197	regulatory protein 10	4.71e+03	904	4	12.9	465	2	S21237	fibronectin K receptor	4.71e+03
832	4	12.9	289	2	G69563	nicotinate-nucleotide	4.71e+03	905	4	12.9	468	1	FG808	fibronogen beta chain	4.71e+03
833	4	12.9	289	1	P8E2A1	phospholipase A1 (EC	4.71e+03	906	4	12.9	471	2	S64310	4-aminobutyrate trans	4.71e+03
834	4	12.9	290	2	JC6012	response-regulator pr	4.71e+03	907	4	12.9	471	2	C71089	hypothetical protein	4.71e+03
835	4	12.9	295	2	S56309	hypothetical protein	4.71e+03	908	4	12.9	475	1	RK1TL	ribulose-bisphosphate	4.71e+03
836	4	12.9	295	2	E64818	probable membrane pro	4.71e+03	909	4	12.9	475	2	S38502	ribulose-bisphosphate	4.71e+03
837	4	12.9	296	2	B28255	hypothetical 34K prot	4.71e+03	910	4	12.9	477	2	S71470	ribulose-bisphosphate	4.71e+03
838	4	12.9	297	2	E64733	nicotinate-nucleotide	4.71e+03	911	4	12.9	478	2	A57174	protein-tyrosine kina	4.71e+03
839	4	12.9	298	2	A64058	dihydrodipicolinate s	4.71e+03	912	4	12.9	478	2	JC4880	fibronolytic metallo	4.71e+03
840	4	12.9	301	2	F69045	imidazoleglycerol-pho	4.71e+03	913	4	12.9	492	2	S62696	catalase (EC 1.11.1.6	4.71e+03
841	4	12.9	306	1	FLB568	flagellin (hag) - Bac	4.71e+03	914	4	12.9	493	2	S71574	mocR protein - Rhizob	4.71e+03
842	4	12.9	306	2	D65008	hypothetical protein	4.71e+03	915	4	12.9	493	2	S77612	amidophosphoribosyltr	4.71e+03
843	4	12.9	309	1	A25776	urate oxidase (EC 1.7	4.71e+03	916	4	12.9	493	2	A38649	1-aminocyclopropane-1	4.71e+03
844	4	12.9	310	2	S20889	superantigen Mtv1 - m	4.71e+03	917	4	12.9	494	2	E70352	anthranilate synthase	4.71e+03
845	4	12.9	312	1	JQ1559	polynhedron envelope p	4.71e+03	918	4	12.9	500	2	S24458	cellulase (EC 3.2.1.4	4.71e+03
846	4	12.9	317	2	JC6147	prostate carcinoma tu	4.71e+03	919	4	12.9	501	2	S18562	xylokainase (EC 2.7.	4.71e+03
847	4	12.9	317	2	S10362	hook-associated prote	4.71e+03	920	4	12.9	505	2	S15193	sucrose porin scry -	4.71e+03
848	4	12.9	320	1	GB0PT4	baseplate protein gp5	4.71e+03	921	4	12.9	509	2	I64183	conserved hypotheticala	4.71e+03
849	4	12.9	322	2	JH0706	superantigen Mtv7 - m	4.71e+03	922	4	12.9	510	2	I64162	mv1N protein homolog	4.71e+03
850	4	12.9	327	2	C55215	2-dehydro-3-deoxygluc	4.71e+03	923	4	12.9	528	1	ALBSK	alpha-amylase (EC 3.2	4.71e+03
851	4	12.9	328	2	A05299	fibronogen beta chain	4.71e+03	924	4	12.9	533	1	TVFVFP	protein-tyrosine kina	4.71e+03
852	4	12.9	329	2	E70973	probable lytB protein	4.71e+03	925	4	12.9	542	2	S47809	lactaldehyde dehydrog	4.71e+03
853	4	12.9	329	2	E70509	probable hemB protei	4.71e+03	926	4	12.9	544	2	S41090	triacylglycerol lipas	4.71e+03
854	4	12.9	329	2	S53833	cytochrome-c oxidase	4.71e+03	927	4	12.9	544	2	S41095	triacylglycerol lipas	4.71e+03
855	4	12.9	334	1	DEBYMM	malate dehydrogenase	4.71e+03	928	4	12.9	551	2	C71606	ATP synthase alpha ch	4.71e+03
856	4	12.9	334	2	C64557	dipeptide ABC transpo	4.71e+03	929	4	12.9	555	2	F65254	ATC transporter in na	4.71e+03
857	4	12.9	335	2	S65043	glyceralddehyde-3-phos	4.71e+03	930	4	12.9	556	2	S13163	t-complex-type molecu	4.71e+03
858	4	12.9	341	1	WMV2H1	antithrombin-III homo	4.71e+03	931	4	12.9	557	2	JT0367	Tcp-1 protein - fruit	4.71e+03
859	4	12.9	342	2	F76626	ABC-type transport pr	4.71e+03	932	4	12.9	563	2	S77533	DNA mismatch repair p	4.71e+03
860	4	12.9	343	2	S25644	Ig mu chain C region	4.71e+03	933	4	12.9	566	1	KIECRU	ribulokinase (EC 2.7.	4.71e+03
861	4	12.9	344	1	BAG5H8	virB11 protein - Agro	4.71e+03	934	4	12.9	572	2	S55982	asparagine synthase (	4.71e+03
862	4	12.9	345	1	WMV2M2	antithrombin-III homo	4.71e+03	935	4	12.9	574	2	S61943	methionine transport	4.71e+03
863	4	12.9	347	2	G00006	haptoglobin - black-h	4.71e+03	936	4	12.9	578	1	BVBFD1	GDH1 protein - yeast	4.71e+03
864	4	12.9	351	2	I64030	hypothetical protein	4.71e+03	937	4	12.9	581	2	S49098	penicillin-binding pr	4.71e+03
865	4	12.9	352	1	HCHU	alpha-1-microglobulin	4.71e+03	938	4	12.9	597	2	E70928	hypothetical protein	4.71e+03
866	4	12.9	364	2	E70449	chaperone DnaJ - Aqu	4.71e+03	939	4	12.9	612	2	S73611	glucose inhibited div	4.71e+03
867	4	12.9	366	2	S70674	bplC protein - Bordet	4.71e+03	940	4	12.9	620	2	E64533	dnak-type molecular c	4.71e+03
868	4	12.9	367	2	I40843	heat shock protein dn	4.71e+03	941	4	12.9	622	2	S56214	probable membrane pro	4.71e+03
869	4	12.9	370	2	D28211	larvicidal toxin 42K	4.71e+03	942	4	12.9	625	2	S71930	neuron-derived recept	4.71e+03
870	4	12.9	370	2	A27098	larvicidal toxin 42K	4.71e+03	943	4	12.9	626	2	S26873	DNA-directed RNA poly	4.71e+03
871	4	12.9	372	1	WMV2CN	major envelope antige	4.71e+03	944	4	12.9	637	2	C65129	hypothetical ABC tran	4.71e+03
872	4	12.9	374	2	F70584	phosphate specific tr	4.71e+03	945	4	12.9	639	2	JC4881	polyvinyl-alcohol deh	4.71e+03
873	4	12.9	376	2	S26435	actin 2 - garden pea	4.71e+03	946	4	12.9	640	1	SVHUAL	5-aminolevulinate syn	4.71e+03
874	4	12.9	380	1	OKBYC2	protein kinase (EC 2.	4.71e+03	947	4	12.9	644	2	S67431	dnak-type molecular c	4.71e+03
875	4	12.9	383	2	C69442	succinyl-coA syntheta	4.71e+03	948	4	12.9	644	2	A64877	exoribonuclease ii (E	4.71e+03
876	4	12.9	384	2	B70675	hypothetical protein	4.71e+03	949	4	12.9	649	2	S36753	dnak-type molecular c	4.71e+03
877	4	12.9	385	2	A47245	acyl-[acyl-carrier-pr	4.71e+03	950	4	12.9	649	2	A33589	mismatch repair prote	4.71e+03
878	4	12.9	394	2	A34417	alpha-1,3-mannosyl-gl	4.71e+03	951	4	12.9	653	1	UZADP2	terminal protein prec	4.71e+03
879	4	12.9	395	1	I50376	homeotic protein lim-	4.71e+03	952	4	12.9	654	1	HHBYS1	dnak-type molecular c	4.71e+03
880	4	12.9	396	2	J70581	hypothetical protein	4.71e+03	953	4	12.9	654	1	HHRTGB	dnak-type molecular c	4.71e+03
881	4	12.9	398	2	QJ0113	2-keto-3-deoxyglucon	4.71e+03	954	4	12.9	656	2	S65826	hypothetical protein	4.71e+03
882	4	12.9	404	2	A36865	ammonium transporter	4.71e+03	955	4	12.9	657	2	A54278	regulatory protein pc	4.71e+03
883	4	12.9	407	2	S30232	histidine--tRNA ligas	4.71e+03	956	4	12.9	659	2	A64139	ribonuclease T2 homol	4.71e+03
884	4	12.9	407	2	A43672	O-antigen polymerase	4.71e+03	957	4	12.9	668	2	A69534	3-hydroxyacyl-CoA deh	4.71e+03
885	4	12.9	412	2	S59068	penton short fiber pr	4.71e+03	958	4	12.9	668	2	D70129	hypothetical protein	4.71e+03
886	4	12.9	413	2	B64983	yeaA protein - Escher	4.71e+03	959	4	12.9	679	1	B48127	dnak-type molecular c	4.71e+03
887	4	12.9	416	2	S39684	ywbN protein - Bacill	4.71e+03	960	4	12.9	679	2	S21764	heat shock protein 82	4.71e+03
888	4	12.9	421	2	C71893	homoserine dehydrogen	4.71e+03	961	4	12.9	688	2	S61249	probable virion prote	4.71e+03
889	4	12.9	421	1	F02M21	M1 protein - vaccinia	4.71e+03	962	4	12.9	716	2	S30687	hypothetical protein	4.71e+03
890	4	12.9	422	1	WMV2M1	E2 protein - bovine p-	4.71e+03	963	4	12.9	735	1	WMBET5	Utl5 protein - human	4.71e+03
891	4	12.9	424	2	B29452	acetyl-coA C-acyltran	4.71e+03	964	4	12.9	735	2	D70174	methyl-accepting chem	4.71e+03
892	4	12.9	430	1	Z4BP33	gene 430 protein - ph	4.71e+03	965	4	12.9	753	2	S59559	prohormone convertase	4.71e+03
893	4	12.9	431	2	S30116	preprotein translocas	4.71e+03	966	4	12.9	750	2	S63138	probable protein kina	4.71e+03
894	4	12.9	437	2	S46613	26S proteasome regula	4.71e+03	967	4	12.9	818	2	JC4397	peroxinectin precurs	4.71e+03
895	4	12.9	438	2	H69992	cytochrome d oxidase	4.71e+03	968	4	12.9	819	2	JC1544	probable leucyl tRNA	4.71e+03
896	4	12.9	441	2	D71191	hypothetical protein	4.71e+03	969	4	12.9	820	1	TVCTFF	protein-tyrosine kina	4.71e+03
897	4	12.9	447	2	S43326	tubulin beta-4 chain	4.71e+03	970	4	12.9	820	2	A40633	chitinase (EC 3.2.1.1	4.71e+03
898	4	12.9	449	2	D69476	TBP-interacting prote	4.71e+03	971	4	12.9	824	2	I50618	c-fps proto oncogene	4.71e+03
899	4	12.9	450	1	DCCHO	ornithine decarboxyla	4.71e+03	972	4	12.9	830	2	A57060	serine/threonine-spec	4.71e+03

```

973      4 12.9 835 2 JC6140      cell surface-associat 4.71e+03
974      4 12.9 837 2 S30971      gene 26 protein - Myc 4.71e+03
975      4 12.9 845 1 JDVLVD      DNA-directed DNA poly 4.71e+03
976      4 12.9 903 1 VGBEK1      glycoprotein B precu 4.71e+03
977      4 12.9 969 2 A70912      probable leus protein 4.71e+03
978      4 12.9 980 2 A38523      genome polyprotein - 4.71e+03
979      4 12.9 1068 2 A43322      phosphatidylinositol 4.71e+03
980      4 12.9 1110 2 I59370      guanylate cyclase (BC 4.71e+03
981      4 12.9 1146 2 I64112      transcription/repair- 4.71e+03
982      4 12.9 1193 2 JC2489      peptidyl-dipeptidase 4.71e+03
983      4 12.9 1196 2 S65245      translation elongatio 4.71e+03
984      4 12.9 1235 1 S16948      insulin receptor subs 4.71e+03
985      4 12.9 1250 2 G44993      yfal protein - Escher 4.71e+03
986      4 12.9 1257 2 A41060      neural cell adhesio 4.71e+03
987      4 12.9 1310 2 I53597      prolone dehydrogenase 4.71e+03
988      4 12.9 1473 2 A33186      salivary agglutinin r 4.71e+03
989      4 12.9 1528 2 A60338      surface antigen A - S 4.71e+03
990      4 12.9 1588 1 BVB1A1      ARO1 protein - yeast 4.71e+03
991      4 12.9 1718 1 JQ1734      genome polyprotein - 4.71e+03
992      4 12.9 1902 2 B44858      serine proteinase (BC 4.71e+03
993      4 12.9 2233 2 S63347      acetyl-CoA carboxylas 4.71e+03
994      4 12.9 2890 2 F64669      DNA-directed RNA poly 4.71e+03
995      4 12.9 3164 1 WMBEH6      UL36 protein - human 4.71e+03
996      4 12.9 3175 1 RRVVEV      genome polyprotein - 4.71e+03
997      4 12.9 3187 2 JC5837      364K Golgi complex-as 4.71e+03
998      4 12.9 5035 2 I45646      ryanodine receptor, s 4.71e+03
999      4 12.9 5037 2 A54161      ryanodine-binding pro 4.71e+03
1000     4 12.9 5147 1 IJFFTM      cadherin-related tumo 4.71e+03

```

## ALIGNMENTS

```

RESULT 1
ENTRY   JG0168      #type complete
TITLE   gob-5 protein - Mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     18-Mar-1999 #sequence_revision 18-Mar-1999 #text_change
18-Mar-1999

ACCESSIONS
REFERENCE JG0168
#authors Komiyu, T.; Tanigawa, Y.; Hirohashi, S.
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351
#title Cloning and identification of the gene gob-5, which is
expressed in intestinal goblet cells in mice.

#accession JG0168
#status preliminary
#residues 1-913 #label KOM
#cross-references DDBJ:AB016592
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

```

```

Query Match 32.3%; Score 10; DB 3; Length 913;
Best Local Similarity 100.0%; Pred. No. 5.67e-07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 645 DNGAGADATK 654
|||||||
QY 1 DNGAGADATK 10

```

```

RESULT 2
ENTRY   A71887      #type complete
TITLE   gmp reductase - Helicobacter pylori (strain J99)
ORGANISM #formal_name Helicobacter pylori
#variety strain J99
DATE     12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
12-Feb-1999

ACCESSIONS
REFERENCE A71887
#authors Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.;
Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge,
B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.;
Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.;
Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis,

```

```

G.F.; Trust, T.J.
Nature (1999) 397:176-180
Genomic sequence comparison of two unrelated isolates of the
human gastric pathogen Helicobacter pylori.
#cross-references MUD:99120557
#accession A71887
#status preliminary
#molecule_type DNA
#residues 1-325 #label ARN
#cross-references GB:AE001509; GB:AE001439; NID:g4155350; PID:g4155370
#experimental_source strain J99
GENETICS
#gene guac
SUMMARY #length 325 #molecular-weight 35845 #checksum 3284
Query Match 22.6%; Score 7; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 159 AGADATK 165
|||||||
QY 4 AGADATK 10

```

```

RESULT 3
ENTRY   C70015      #type complete
TITLE   GMP reductase (EC 1.6.6.8) homolog yumd - Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE     05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
24-Sep-1998

```

```

ACCESSIONS
REFERENCE C70015
#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Prescean, E.; Pujić, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;
Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumsteln, E.;
Yoshikawa, H.; Danchin, A.
Nature (1997) 390:249-256
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references MUD:9804033
#accession C70015
#status preliminary; nucleic acid sequence not shown;

```

```

translation not shown
##molecule_type DNA
##residues 1-326 ##label KUN
##cross-references GB:299120; GB:AL009126; NID:g2635613; PID:e1184292;
##experimental_source strain 168
GENETICS
#gene
#keywords
#summary
yumd
NAD: oxidoreductase
#length 326 #molecular-weight 35819 #checksum 7996
Query Match 22.6%; Score 7; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 160 AGADATK 166
|||||
Qy 4 AGADATK 10
RESULT 4
ENTRY F64626 #type complete
TITLE GMP reductase [EC 1.5.6.8] homolog - Helicobacter pylori
ORGANISM (strain 26695)
#formal_name Helicobacter pylori
DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change
22-May-1998
F64626
A64520
ACCESSIONS
REFERENCE Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.;
Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk,
H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush,
J.P.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;
Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.;
McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.;
Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.;
Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.;
Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes,
W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,
C.M.; Venter, J.C.
#journal Nature (1997) 388:539-547
#title The complete genome sequence of the gastric pathogen
Helicobacter pylori.
#cross-references MUID:97394467
#accession F64626
##status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-327 ##label TOM
##cross-references GB:AE000596; GB:AE000511; NID:g2313982; PID:g2313987;
TIGR:HP0854
KEYWORDS NADP: oxidoreductase
SUMMARY #length 327 #molecular-weight 36038 #checksum 7430
Query Match 22.6%; Score 7; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 161 AGADATK 167
|||||
Qy 4 AGADATK 10
RESULT 5
ENTRY I40300 #type complete
TITLE outer membrane lipoprotein - Borrelia hermsii
ORGANISM #formal_name Borrelia hermsii
DATE 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change
12-Aug-1996
I40300
ACCESSIONS
REFERENCE Restrepo, B.I.; Kitten, T.; Carter, C.J.; Infante, D.;
Barbour, A.G.
#journal Mol. Microbiol. (1992) 6:3299-3311
#title Subtelomeric expression regions of Borrelia hermsii linear
plasmids are highly polymorphic.
#cross-references MUID:93133110
#accession I40300
##status preliminary; translated from GB/EMBL/DBJ
##molecule_type DNA
##residues 1-353 ##label RES
##cross-references GB:L04788; NID:g144023; PID:g144024
GENETICS
#gene vmp17
#summary #length 353 #molecular-weight 36460 #checksum 8369
Query Match 22.6%; Score 7; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 170 ADATKDD 176
|||||
Qy 6 ADATKDD 12
RESULT 6
ENTRY S41689 #type fragment
TITLE cytochrome-c oxidase (EC 1.9.3.1) chain I - Theileria parva
mitochondrion (SGC6) (fragment)
ORGANISM #formal_name mitochondrion Theileria parva
DATE 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change
05-Dec-1998
S41689; S40154
ACCESSIONS
REFERENCE S41689
#authors Kaito, A.; Fairlamb, A.H.; Gobright, E.; Nene, V.
#journal EMBO J. (1994) 13:898-905
#title A 7.1 kb linear DNA molecule of Theileria parva has scrambled
rDNA sequences and open reading frames for mitochondrially
encoded proteins.
#cross-references MUID:94153854
#accession S41689
##status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-483 ##label KAI
##cross-references EMBL:223263; NID:g437862; PID:g437863
##note the nucleotide sequence was submitted to the EMBL Data
Library, July 1993
GENETICS
#genome mitochondrion
#genetic_code SGC6
CLASSIFICATION #superfamily cytochrome-c oxidase chain I; cytochrome-c
oxidase chain I homology
KEYWORDS chromoprotein; copper binding; electron transfer; heme;
membrane-associated complex; mitochondrion; oxidative
phosphorylation; oxidoreductase; respiratory chain;
transmembrane protein
FEATURE
18-466 #domain cytochrome-c oxidase chain I homology #label
COI\
71,366 #binding_site heme a iron (His) (axial ligands) #status
predicted\
384 #binding_site heme a3 iron (His) (axial ligand) #status
predicted
SUMMARY #length 483 #checksum 8053
Query Match 22.6%; Score 7; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 308 YSRYFTT 314
|||||
Qy 15 YSRYFTT 21
RESULT 7
ENTRY S64942 #type complete

```

```

TITLE      probable membrane protein YLR106c - yeast (Saccharomyces
ALTERNATE_NAMES  cerevisiae)
ORGANISM    hypothetical protein L2901
DATE        #formal_name Saccharomyces cerevisiae
            #01-Aug-1995 #sequence_revision 24-May-1996 #text_change
            21-Nov-1997
ACCESSIONS  S64942; S64943; S69393
REFERENCE    S64935
            #authors      Messenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.
            #submission   submitted to the Protein Sequence Database, May 1996
            #accession    S64942
            #molecule_type DNA
            ##residues    1-4910 #label MES
            ##cross-references EMBL:Z73278; NID:g1360493; PID:e245804; PID:g1360494;
            MIPS:YLR106c
            ##experimental_source strain S288C
REFERENCE    S64943
            #authors      Verhasselt, P.; Voet, M.; Volckaert, G.
            #submission   submitted to the Protein Sequence Database, May 1996
            #accession    S64943
            #molecule_type DNA
            ##residues    1-115 #label VPR
            ##cross-references EMBL:Z73278; MIPS:YLR106c
            ##experimental_source strain S288C
REFERENCE    S69393
            #authors      Verhasselt, P.; Volckaert, G.
            #submission   submitted to the EMBL Data Library, September 1995
            #accession    S69393
            #molecule_type DNA
            ##residues    1-115 #label VW
            ##cross-references EMBL:X89514; NID:g1297019; PID:e198739; PID:g1297020
GENETICS
            #map_position 12R
            #transmembrane protein
FEATURE      85-102      #domain transmembrane #status predicted #label TM1\
            174-190      #domain transmembrane #status predicted #label TM2\
            1879-1895     #domain transmembrane #status predicted #label TM3\
            2512-2528     #domain transmembrane #status predicted #label TM4\
            2547-2563     #domain transmembrane #status predicted #label TM5\
            2893-2909     #domain transmembrane #status predicted #label TM6\
SUMMARY      #length 4910 #molecular-weight 559302 #checksum 5145
Query Match      22.6%; Score 7; DB 2; Length 4910;
Best Local Similarity 100.0%; Pred. No. 2.12e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 4405 GAGADAT 4411
Qy 3 GAGADAT 9
            |||||

RESULT 8
ENTRY   A33717      #type fragments
TITLE   nitrate reductase (nicotinamide-dependent) (EC 1.6.6.-) -
ORGANISM Chlamydomonas reinhardtii (fragments)
DATE     27-Feb-1990 #sequence_revision 27-Feb-1990 #text_change
            07-Aug-1998
ACCESSIONS A33717
REFERENCE   A33717
            #authors      Fernandez, E.; Schnell, R.; Ranum, L.P.W.; Hussey, S.C.;
            Silflow, C.D.; Lefebvre, P.A.
            #journal      Proc. Natl. Acad. Sci. U.S.A. (1989) 86:6449-6453
            #title        Isolation and characterization of the nitrate reductase
            structural gene of Chlamydomonas reinhardtii.
            #cross-references MUID:89367266
            #accession    A33717
            ##status      preliminary
            ##molecule_type DNA
            ##residues    1-62 #label FER
            ##cross-references GB:M26074; GB:M26075
CLASSIFICATION #superfamily nitrate reductase (NADH); cytochrome b5 core

```

```

homology; cytochrome-b5 reductase homology;
molybdopterin-binding domain homology
dimer; electron transfer; FAD; flavoprotein; heme;
metalloprotein; molybdenum; nitrate assimilation;
oxidoreductase
#domain cytochrome b5 core homology (fragment) #label
CB5\
#domain cytochrome-b5 reductase homology (fragment)
#label CBR
#length 62 #checksum 2312
FEATURE      1-23
            23-62
SUMMARY      #length 62 #checksum 2312
Query Match      19.4%; Score 6; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 11 AGADAT 16
Qy 4 AGADAT 9
            |||||

RESULT 9
ENTRY   S36319      #type fragment
TITLE   T-cell receptor delta chain precursor - sheep (fragment)
ORGANISM #formal_name Ovis orientalis aries, Ovis ammon aries
            #common_name domestic sheep
DATE     03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
            17-Mar-1999
ACCESSIONS S36319; S23034
REFERENCE   S36287
            #authors      Hein, W.R.; Dudler, L.
            #journal      EMBO J. (1993) 12:715-724
            #title        Divergent evolution of T cell repertoires: extensive
            diversity and developmentally regulated expression of the
            sheep gamma-delta T cell receptor.
            #cross-references MUID:93178447
            #accession    S36319
            ##status      preliminary; translation not shown
            ##molecule_type mRNA
            ##residues    1-87 #label HEI
            ##cross-references EMBL:Z12991; NID:g2246; PID:g2247
CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS      T-cell receptor
SUMMARY      #length 87 #checksum 534
Query Match      19.4%; Score 6; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 80 NGRYSV 85
Qy 25 NGRYSV 30
            |||||

RESULT 10
ENTRY   S07110      #type complete
TITLE   T-cell receptor delta chain V region (253) - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE     21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
            10-Nov-1995
ACCESSIONS S07110
REFERENCE   S06342
            #authors      Elliott, J.F.; Rock, E.P.; Patten, P.A.; Davis, M.M.; Chien,
            Y.H.
            #journal      Nature (1988) 331:627-631
            #title        The adult T-cell receptor delta-chain is diverse and distinct
            from that of fetal thymocytes.
            #cross-references MUID:88122668
            #accession    S07110
            ##status      preliminary
            ##molecule_type mRNA
            ##residues    1-95 #label ELL
CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology

```

```
KEYWORDS      T-cell receptor
SUMMARY       #length 95 #molecular-weight 10925 #checksum 6256

Query Match   19.4%; Score 6; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 61 NGRYSV 66
QY 25 NGRYSV 30

RESULT 11
ENTRY   I40736 #type complete
TITLE   ItuB protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)
ORGANISM #formal_name Chlamydia trachomatis
DATE     16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change
ACCESSIONS I40736; B71560
REFERENCE  #authors
#journal   Fahr, M.J.; Douglas, A.L.; Xia, W.; Hatch, T.P.
#title     J. Bacteriol. (1995) 177:4252-4260
#title     Characterization of late gene promoters of Chlamydia trachomatis.
#cross-references MUID:95362647
#accession I40736
#status     preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-97 #label RES
#cross-references GB:I40838; NID:g886219; PID:g886220
REFERENCE  A71570
#authors    Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, O.; Koonin, E.V.; Davis, R.W.
#journal    Science (1998) 282:754-759
#title      Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.
#cross-references MUID:9900809
#accession B71560
#molecule_type DNA
#residues 1-97 #label ARN
#cross-references GB:AE001282; GB:AE001273; NID:g3328466; PID:g3328475
#experimental_source serotype D, strain UW-3/Cx

GENETICS
#gene       ituB
SUMMARY     #length 97 #molecular-weight 11323 #checksum 8359

Query Match   19.4%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 91 GRYSVK 96
QY 26 GRYSVK 31

RESULT 12
ENTRY   S36305 #type fragment
TITLE   T-cell receptor delta chain - sheep (fragment)
ORGANISM #formal_name Ovis orientalis aries, Ovis ammon aries
#common_name domestic sheep
DATE     03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
ACCESSIONS S36305; S23021
REFERENCE  #authors
#journal   Hein, W.R.; Dudley, L.
#title     EMBO J. (1993) 12:715-724
#title     Divergent evolution of T cell repertoires: extensive diversity and developmentally regulated expression of the sheep gamma-delta T cell receptor.
#cross-references MUID:93178447
#accession S36305

#status     preliminary; translation not shown
#molecule_type mRNA
#residues 1-104 #label HEI
#cross-references EMBL:212977; NID:g2219; PID:g2220
CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS      T-cell receptor
SUMMARY       #length 104 #checksum 5378

Query Match   19.4%; Score 6; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 NGRYSV 47
QY 25 NGRYSV 30

RESULT 13
ENTRY   S36306 #type fragment
TITLE   T-cell receptor delta chain - sheep (fragment)
ORGANISM #formal_name Ovis orientalis aries, Ovis ammon aries
#common_name domestic sheep
DATE     03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
ACCESSIONS S36306; S23022
REFERENCE  #authors
#journal   Hein, W.R.; Dudley, L.
#title     EMBO J. (1993) 12:715-724
#title     Divergent evolution of T cell repertoires: extensive diversity and developmentally regulated expression of the sheep gamma-delta T cell receptor.
#cross-references MUID:93178447
#accession S36306
#status     preliminary; translation not shown
#molecule_type mRNA
#residues 1-120 #label HEI
#cross-references EMBL:212978; NID:g2221; PID:g2222
CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS      T-cell receptor
SUMMARY       #length 120 #checksum 6644

Query Match   19.4%; Score 6; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 NGRYSV 61
QY 25 NGRYSV 30

RESULT 14
ENTRY   CBRT5M #type complete
TITLE   Cytochrome b5, outer mitochondrial membrane - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE     04-Dec-1986 #sequence_revision 29-Aug-1997 #text_change
ACCESSIONS A00173; S66501
REFERENCE  #authors
#journal   Lederer, F.; Ghir, R.; Guiard, B.; Cortial, S.; Ito, A.
#title     Eur. J. Biochem. (1983) 132:95-102
#title     Two homologous cytochromes b-5 in a single cell.
#cross-references MUID:83182449
#accession A00173
#molecule_type protein
#residues 1-92 #label LED
REFERENCE  S66501
#authors    de Silvestris, M.; D'Arrigo, A.; Borgese, N.
#journal    FEBS Lett. (1995) 370:69-74
#title      The targeting information of the mitochondrial outer membrane isoform of cytochrome b(5) is contained within the carboxyl-terminal region.
#cross-references MUID:95377460
#accession S66501
```

```
##molecule_type mRNA
##residues 32-135 #label DES
##cross-references GB:X96392; EMBL:S79339; NID:gi217654; PID:e225547;

COMMENT Cytochrome b5, found attached to various hepatic cell membranes, is
a major component of the electron transport system, catalyzing
the NADH-linked desaturation of fatty acids in the endoplasmic
reticulum. It may also be involved in the NADH-linked pathway of
drug hydroxylation reactions catalyzed by cytochrome p450.

GENETICS
#genome nuclear
#superfamily cytochrome b5; cytochrome b5 core homology
#chromoprotein; electron transfer; heme; iron; membrane
#protein; mitochondrion
#domain heme binding #status predicted #label HMB\
#domain cytochrome b5 core homology #label CBS\
#binding_site heme iron (His) (axial ligands) #status
predicted
SUMMARY #length 135 #molecular-weight 15221 #checksum 281

Query Match 19.4%; Score 6; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 AGADAT 60
Qy 4 AGADAT 9
|||||

RESULT 15
ENTRY S36320 #type fragments
TITLE T-cell receptor delta chain precursor - sheep (fragments)
ORGANISM #formal_name Ovis orientalis aries, Ovis ammon aries
#common_name domestic sheep
DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
17-Mar-1999
ACCESSIONS S36320; S23035
REFERENCE S36287
#authors Hein, W.R.; Dudler, L.
#journal EMBO J. (1993) 12:715-724
#title Divergent evolution of T cell repertoires: extensive
diversity and developmentally regulated expression of the
sheep gamma-delta T cell receptor.
#cross-references MUID:93178447
#accession S36320
#status preliminary; translation not shown
##molecule_type mRNA
##residues 1-136 #label HEI
##cross-references EMBL:212992; NID:g2248; PID:g2249
CLASSIFICATION #superfamily immunoglobulin V region; immunoglobulin homology
KEYWORDS T-cell receptor
SUMMARY #length 136 #checksum 7888

Query Match 19.4%; Score 6; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 9.25e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 NGRYSV 85
Qy 25 NGRYSV 30
|||||

Search completed: Sat Aug 28 15:03:38 1999
Job time : 44 secs.
```

\*\*\*\*\*

WISN (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 15:00:47 1999; MasPar time 3.46 Seconds  
Tabular output not generated.

Title: >US-09-049-696-47  
Description: (1-31) from US09049696.pep  
Sequence: 1 DNGAGADATKDGVSRYFTYDTNGRYSVK 31  
Scoring table: TABLE unitprotatable  
Gap 60

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 2.779; Variance 0.421; scale 6.598

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	7	22.6	353	1	VM17_BORHE VARIABLE MAJOR OUTER M	8.25e-02
2	7	22.6	903	1	ECIC_BOVIN EPITHELIAL CHLORIDE CH	8.25e-02
3	6	19.4	92	1	CYM5_RAT CYTOCHROME B5, OUTER M	4.59e+00
4	6	19.4	279	1	ATND_RAT SODIUM/POTASSIUM-TRANS	4.59e+00
5	6	19.4	295	1	DAPP_METJA DIAMINOPIMELATE EPIMER	4.59e+00
6	6	19.4	327	1	GSPK_ECOLI PROBABLE GENERAL SECURE	4.59e+00
7	6	19.4	328	1	Y051_CAEEL PUTATIVE MITOCHONDRIAL	4.59e+00
8	6	19.4	372	1	SAXC_ECOLI PUTATIVE SARGOSINE OXI	4.59e+00
9	6	19.4	378	1	CARA_PSEAE CARBAMOYL-PHOSPHATE SY	4.59e+00
10	6	19.4	384	1	CARA_PSEST CARBAMOYL-PHOSPHATE SY	4.59e+00
11	6	19.4	387	1	COAT_TBSVB COAT PROTEIN.	4.59e+00
12	6	19.4	514	1	AMY_BACAM ALPHA-AMYLASE PRECURSOR	4.59e+00
13	6	19.4	701	1	SECD_MYCLE PROTEIN-EXPORT MEMBRAN	4.59e+00
14	6	19.4	747	1	FH0A_ECOLI FERRICHYROM-IRON RECEP	4.59e+00
15	6	19.4	799	1	FLO8_YEAST TRANSCRIPTIONAL ACTIVA	4.59e+00
16	6	19.4	864	1	NIA_VOLCA NITRATE REDUCTASE (EC	4.59e+00
17	6	19.4	1131	1	YAB9_YEAST HYPOTHETICAL 128.5 KD	4.59e+00
18	6	19.4	1426	1	TOP_DROME GURKEN RECEPTOR PRECUR	4.59e+00
19	6	19.4	1444	1	ADP1_MYCGE ADHESIN P1 PRECURSOR (	4.59e+00
20	5	16.1	37	1	PSAJ_EUGGR PHOTOSYSTEM I REACTION	1.65e+02
21	5	16.1	44	1	PHI1_CHRSP PHYCOERYTHRIN ALPHA-1	1.65e+02
22	5	16.1	52	1	ALIA3_HORSE ALPHA-1-ANTIPROTEINASE	1.65e+02
23	5	16.1	53	1	ALIA1_HORSE ALPHA-1-ANTIPROTEINASE	1.65e+02





243	5	16.1	437	1	NCCC_ALCXX	1.65e+02	316	5	16.1	567	1	G6P1_ORYSA	GLUCOSE-6-PHOSPHATE IS	1.65e+02
244	5	16.1	442	1	G6P1_MYCSM	1.65e+02	317	5	16.1	567	1	G6P1_MAIZE	GLUCOSE-6-PHOSPHATE IS	1.65e+02
245	5	16.1	445	1	G6PB_BACST	1.65e+02	318	5	16.1	568	1	G6P1_CLAAR	GLUCOSE-6-PHOSPHATE IS	1.65e+02
246	5	16.1	446	1	CIIN_SALPU	1.65e+02	319	5	16.1	568	1	G6P1_CLAWI	GLUCOSE-6-PHOSPHATE IS	1.65e+02
247	5	16.1	446	1	CIIN_SALDU	1.65e+02	320	5	16.1	568	1	G6P1_CLAMI	GLUCOSE-6-PHOSPHATE IS	1.65e+02
248	5	16.1	446	1	CIIN_KLEPN	1.65e+02	321	5	16.1	568	1	G6PB_ORYSA	GLUCOSE-6-PHOSPHATE IS	1.65e+02
249	5	16.1	448	1	PRTE_BACNO	1.65e+02	322	5	16.1	568	1	G6P1_CLAFR	GLUCOSE-6-PHOSPHATE IS	1.65e+02
250	5	16.1	449	1	G6PA_BACST	1.65e+02	323	5	16.1	568	1	G6P1_CLAXA	GLUCOSE-6-PHOSPHATE IS	1.65e+02
251	5	16.1	450	1	BEDI_PSEPU	1.65e+02	324	5	16.1	568	1	G6P1_OENME	GLUCOSE-6-PHOSPHATE IS	1.65e+02
252	5	16.1	450	1	G6P1_BACSU	1.65e+02	325	5	16.1	569	1	G6P1_CLALE	GLUCOSE-6-PHOSPHATE IS	1.65e+02
253	5	16.1	454	1	IMDH_PNECA	1.65e+02	326	5	16.1	569	1	G6P2_CLAXA	GLUCOSE-6-PHOSPHATE IS	1.65e+02
254	5	16.1	457	1	TRPE_THEMEA	1.65e+02	327	5	16.1	569	1	G6P1_CLACO	GLUCOSE-6-PHOSPHATE IS	1.65e+02
255	5	16.1	458	1	ACHO_HUMAN	1.65e+02	328	5	16.1	569	1	G6P2_CLACO	GLUCOSE-6-PHOSPHATE IS	1.65e+02
256	5	16.1	458	1	HN3B_RAT	1.65e+02	329	5	16.1	569	1	G6P2_CLALE	GLUCOSE-6-PHOSPHATE IS	1.65e+02
257	5	16.1	459	1	HN3B_MOUSE	1.65e+02	330	5	16.1	570	1	G6P1_CLARO	HEMAGGLUTININ-NEURAMIN	1.65e+02
258	5	16.1	461	1	Y4SH_RHISN	1.65e+02	331	5	16.1	570	1	HEMA_NDVA	CELL SURFACE ANTIGEN 1	1.65e+02
259	5	16.1	461	1	Y4UA_RHISN	1.65e+02	332	5	16.1	573	1	C114_MOUSE	ENV POLYPROTEIN PRECUR	1.65e+02
260	5	16.1	466	1	HN3A_RAT	1.65e+02	333	5	16.1	575	1	ENV_SMRVH	HEMAGGLUTININ-NEURAMIN	1.65e+02
261	5	16.1	468	1	HN3A_MOUSE	1.65e+02	334	5	16.1	577	1	HEMA_NDVH4	HEMAGGLUTININ-NEURAMIN	1.65e+02
262	5	16.1	473	1	HN3A_HUMAN	1.65e+02	335	5	16.1	577	1	HEMA_NDVH	HEMAGGLUTININ-NEURAMIN	1.65e+02
263	5	16.1	473	1	GUNB_NEOPA	1.65e+02	336	5	16.1	585	1	PTRR_DIDMA	PARATHYROID HORMONE/PA	1.65e+02
264	5	16.1	474	1	TRPC_BRELA	1.65e+02	337	5	16.1	590	1	VP02_BPPRD	ADSORPTION PROTEIN P2.	1.65e+02
265	5	16.1	474	1	SOXA_HUMAN	1.65e+02	338	5	16.1	593	1	DCEL_MOUSE	GLUTAMATE DECARBOXYLAS	1.65e+02
266	5	16.1	479	1	PK2_DICDI	1.65e+02	339	5	16.1	593	1	DCEL_RAT	GLUTAMATE DECARBOXYLAS	1.65e+02
267	5	16.1	479	1	PRTC_ERMCH	1.65e+02	340	5	16.1	594	1	DCEL_FELCA	GLUTAMATE DECARBOXYLAS	1.65e+02
268	5	16.1	479	1	APRA_FSEAE	1.65e+02	341	5	16.1	594	1	DCEL_HUMAN	GLUTAMATE DECARBOXYLAS	1.65e+02
269	5	16.1	481	1	IMDH_HELPY	1.65e+02	342	5	16.1	594	1	DCEL_PIG	GLUTAMATE DECARBOXYLAS	1.65e+02
270	5	16.1	484	1	CATA_CANAL	1.65e+02	343	5	16.1	595	1	TRPE_ARATH	ANTHRANILATE SYNTHASE	1.65e+02
271	5	16.1	485	1	IMDH_PYRFO	1.65e+02	344	5	16.1	601	1	DNJM_MYCGE	DNAA-LIKE PROTEIN MG20	1.65e+02
272	5	16.1	488	1	IMDH_ACICA	1.65e+02	345	5	16.1	605	1	G6P1_LEIME	GLUCOSE-6-PHOSPHATE IS	1.65e+02
273	5	16.1	488	1	CATA_LISSE	1.65e+02	346	5	16.1	607	1	G6P1_TRYBB	GLUCOSE-6-PHOSPHATE IS	1.65e+02
274	5	16.1	488	1	ASTD_FSEAE	1.65e+02	347	5	16.1	612	1	PUBL_PENNO	LYSOPHOSPHOLIPASE PREC	1.65e+02
275	5	16.1	489	1	VLGC_HSVTH	1.65e+02	348	5	16.1	613	1	PKPA_PHYBL	SERINE/THREONINE PROTE	1.65e+02
276	5	16.1	490	1	YPLC_CLOPE	1.65e+02	349	5	16.1	615	1	AIPI_YEAST	ACTIN INTERACTING PROT	1.65e+02
277	5	16.1	493	1	IMDH_STRPY	1.65e+02	350	5	16.1	616	1	HEMA_NDVU	HEMAGGLUTININ-NEURAMIN	1.65e+02
278	5	16.1	496	1	IMDH_METJA	1.65e+02	351	5	16.1	616	1	HEMA_NDVD	HEMAGGLUTININ-NEURAMIN	1.65e+02
279	5	16.1	496	1	CHK1_SCHPO	1.65e+02	352	5	16.1	616	1	HEMA_NDVQ	HEMAGGLUTININ-NEURAMIN	1.65e+02
280	5	16.1	497	1	MEK1_YEAST	1.65e+02	353	5	16.1	618	1	XIA2_BACST	BETA-XILOSIDASE PRECUR	1.65e+02
281	5	16.1	497	1	VLI_HPV54	1.65e+02	354	5	16.1	619	1	HEX3_YEAST	HEXOSE METABOLISM-RELA	1.65e+02
282	5	16.1	498	1	YCKJ_ECOLI	1.65e+02	355	5	16.1	621	1	Y04E_MYCTU	HYPOTHETICAL 69.2 KD A	1.65e+02
283	5	16.1	500	1	TRPE_RHOSH	1.65e+02	356	5	16.1	621	1	TRPX_ARATH	ANTHRANILATE SYNTHASE	1.65e+02
284	5	16.1	503	1	CD44_RAT	1.65e+02	357	5	16.1	627	1	FLGK_BORBU	FLAGELLAR HOOK-ASSOCIA	1.65e+02
285	5	16.1	504	1	CHIT_BRUMA	1.65e+02	358	5	16.1	628	1	Y0D8_MYCTU	HYPOTHETICAL 69.9 KD P	1.65e+02
286	5	16.1	507	1	YIK4_YEAST	1.65e+02	359	5	16.1	630	1	YD13_SCHPO	HYPOTHETICAL 70.6 KD P	1.65e+02
287	5	16.1	510	1	ATPA_RHOUR	1.65e+02	360	5	16.1	634	1	DNAA_HAEIN	DNAA PROTEIN (HEAT SHO	1.65e+02
288	5	16.1	512	1	AMY_BACLI	1.65e+02	361	5	16.1	637	1	DNAA_ECOLI	DNAA PROTEIN (HEAT SHO	1.65e+02
289	5	16.1	515	1	EFIS_PORPU	1.65e+02	362	5	16.1	641	1	HEMA_INCA	HEMAGGLUTININ	1.65e+02
290	5	16.1	518	1	ANT6_BACST	1.65e+02	363	5	16.1	641	1	HEMA_INCHY	HEMAGGLUTININ	1.65e+02
291	5	16.1	519	1	LAC2_TRAVI	1.65e+02	364	5	16.1	641	1	HEMA_INCM1	HEMAGGLUTININ PRECURSO	1.65e+02
292	5	16.1	519	1	LAC2_TRAVE	1.65e+02	365	5	16.1	641	1	HEMA_INCKY	HEMAGGLUTININ PRECURSO	1.65e+02
293	5	16.1	520	1	LAC1_CORHI	1.65e+02	366	5	16.1	642	1	HEMA_INCP2	HEMAGGLUTININ PRECURSO	1.65e+02
294	5	16.1	520	1	LAC1_TRAVI	1.65e+02	367	5	16.1	642	1	HEMA_INCP3	HEMAGGLUTININ PRECURSO	1.65e+02
295	5	16.1	521	1	IMDH_CHLVI	1.65e+02	368	5	16.1	642	1	HEMA_INCG1	HEMAGGLUTININ PRECURSO	1.65e+02
296	5	16.1	521	1	IMH3_CANAL	1.65e+02	369	5	16.1	642	1	HEMA_INCN1	HEMAGGLUTININ PRECURSO	1.65e+02
297	5	16.1	523	1	UL21_PRVN3	1.65e+02	370	5	16.1	642	1	HEMA_INCN3	HEMAGGLUTININ PRECURSO	1.65e+02
298	5	16.1	529	1	IMDH_MYCLE	1.65e+02	371	5	16.1	642	1	HEMA_INCTA	HEMAGGLUTININ PRECURSO	1.65e+02
299	5	16.1	529	1	IMDH_MYCTU	1.65e+02	372	5	16.1	642	1	HEMA_INCP3	HEMAGGLUTININ PRECURSO	1.65e+02
300	5	16.1	530	1	UTR1_YEAST	1.65e+02	373	5	16.1	648	1	TALA_POVMK	LARGE T ANTIGEN.	1.65e+02
301	5	16.1	530	1	UDBE_RABIT	1.65e+02	374	5	16.1	654	1	HEMA_INCCA	HEMAGGLUTININ PRECURSO	1.65e+02
302	5	16.1	535	1	ARSB_FELCA	1.65e+02	375	5	16.1	655	1	CD44_MOUSE	CD44 ANTIGEN PRECURSOR	1.65e+02
303	5	16.1	536	1	DITL_YEAST	1.65e+02	376	5	16.1	655	1	HEMA_INCUH	HEMAGGLUTININ PRECURSO	1.65e+02
304	5	16.1	538	1	C0X1_RHILE	1.65e+02	377	5	16.1	658	1	HRSA_ECOLI	HRSA PROTEIN (EC 2.7.1	1.65e+02
305	5	16.1	545	1	CH60_PSEPU	1.65e+02	378	5	16.1	661	1	YK65_CAEEL	HYPOTHETICAL 73.3 KD P	1.65e+02
306	5	16.1	548	1	TLRC_STRFR	1.65e+02	379	5	16.1	668	1	MTMW_METWO	MODIFICATION METHYLASE	1.65e+02
307	5	16.1	550	1	G6P1_SCHPO	1.65e+02	380	5	16.1	672	1	COP1_PEA	COP1 REGULATORY PROTEI	1.65e+02
308	5	16.1	553	1	G6P1_YEAST	1.65e+02	381	5	16.1	673	1	YD13_AGRF5	COP1 REGULATORY PROTEI	1.65e+02
309	5	16.1	553	1	G6P1_MYCTU	1.65e+02	382	5	16.1	675	1	COP1_ARATH	COP1 REGULATORY PROTEI	1.65e+02
310	5	16.1	555	1	OL6G_BACCO	1.65e+02	383	5	16.1	677	1	GYBR_STRSH	DNA GYRASE SUBUNIT B,	1.65e+02
311	5	16.1	557	1	V1FP_BPT3	1.65e+02	384	5	16.1	678	1	NCPR_SCHPO	NADPH-CYTOCHROME P450	1.65e+02
312	5	16.1	557	1	G6P1_KLULA	1.65e+02	385	5	16.1	698	1	CVAB_ECOLI	COLICIN V SECRETION AT	1.65e+02
313	5	16.1	560	1	G6P1_ARATH	1.65e+02	386	5	16.1	704	1	YBY1_YEAST	HYPOTHETICAL 80.7 KD P	1.65e+02
314	5	16.1	566	1	PRIM_BPT7	1.65e+02	387	5	16.1	713	1	SG2N_HUMAN	CELL-CYCLE NUCLEAR AUT	1.65e+02
315	5	16.1	566	1	PRIM_BPT3	1.65e+02	388	5	16.1	714	1	YUL2_YEAST	HYPOTHETICAL 80.0 KD T	1.65e+02

389	5	16.1	720	1	KRE6_YEAST	BETA-GLUCAN SYNTHESIS-	1.65e+02	462	1	RT32_ACTPL	RTX-III TOXIN DETERMIN	1.65e+02
390	5	16.1	728	1	CATB_ASPTU	CATALASE B (EC 1.11.1.1-	1.65e+02	463	1	GUNC_CELFI	ENDOGLUCANASE C PRECUR	1.65e+02
391	5	16.1	729	1	FHUE_ECOLI	OUTER-MEMBRANE RECEPTO	1.65e+02	464	1	COLA_CLOPE	MICROBIAL COLLAGENASE	1.65e+02
392	5	16.1	730	1	DCOR_LACS3	ORNITHINE DECARBOXYLAS	1.65e+02	465	1	SYLC_SCHPO	PUTATIVE LEUCYL-TRNA S	1.65e+02
393	5	16.1	736	1	VML_REOVL	MINOR VIRION STRUCTURA	1.65e+02	466	1	Y140_MYCGE	HYPOTHETICAL ATP-BINDI	1.65e+02
394	5	16.1	736	1	VML_REOVL	MINOR VIRION STRUCTURA	1.65e+02	467	1	RPB2_DROME	DNA-DIRECTED RNA POLYM	1.65e+02
395	5	16.1	736	1	MYSB_RABIT	MYOSIN HEAVY CHAIN, CA	1.65e+02	468	1	YB95_YEAST	HYPOTHETICAL 131.1 KD	1.65e+02
396	5	16.1	737	1	SKN1_CANAL	BETA-GLUCAN SYNTHESIS-	1.65e+02	469	1	POLN_SHSV1	NON-STRUCTURAL POLYPRO	1.65e+02
397	5	16.1	740	1	KRE6_CANAL	BETA-GLUCAN SYNTHESIS-	1.65e+02	470	1	RPB2_HUMAN	DNA-DIRECTED RNA POLYM	1.65e+02
398	5	16.1	742	1	CD44_HUMAN	CD44 ANTIGEN PRECURSOR	1.65e+02	471	1	NIR_NEUCR	NITRITE REDUCTASE (NAD	1.65e+02
399	5	16.1	744	1	MUTA_CAEEL	PROBABLE METHYLMALONYL	1.65e+02	472	1	METH_MYCTU	5-METHYLTETRAHYDROFOLA	1.65e+02
400	5	16.1	745	1	CYAA_TRYCO	RECEPTOR-TYPE ADENYLAT	1.65e+02	473	1	RPB2_CAEEL	DNA-DIRECTED RNA POLYM	1.65e+02
401	5	16.1	750	1	MUTA_MYCTU	METHYLMALONYL-COA MUTA	1.65e+02	474	1	EVGS_ECOLI	PUTATIVE SENSOR PROTEI	1.65e+02
402	5	16.1	752	1	CTPB_MYCTU	CATION-TRANSPORTING P-	1.65e+02	475	1	IRBP_HUMAN	INTERPHOTORECEPTOR RET	1.65e+02
403	5	16.1	752	1	CO2_HUMAN	COMPLEMENT C2 PRECURSO	1.65e+02	476	1	YKV5_YEAST	HYPOTHETICAL 140.4 KD	1.65e+02
404	5	16.1	756	1	MHL1_HUMAN	MUTL PROTEIN HOMOLOG 1	1.65e+02	477	1	VAC2_HELPY	VACUOLATING CYTOTOXIN	1.65e+02
405	5	16.1	757	1	HT16_HYDAT	TYROSINE-PROTEIN KINAS	1.65e+02	478	1	VACO_HELPY	VACUOLATING CYTOTOXIN	1.65e+02
406	5	16.1	764	1	CFAB_HUMAN	COMPLEMENT FACTOR B PR	1.65e+02	479	1	VAC1_HELPY	VACUOLATING CYTOTOXIN	1.65e+02
407	5	16.1	768	1	TRPG_COCH	ANTHRANILATE SYNTHASE	1.65e+02	480	1	VAC1_HELPY	VACUOLATING CYTOTOXIN	1.65e+02
408	5	16.1	771	1	SRN1_YEAST	BETA-GLUCAN SYNTHESIS-	1.65e+02	481	1	ICP4_HSV11	TRANS-ACTING TRANSCRIP	1.65e+02
409	5	16.1	774	1	RRP3_INCB	RNA-DIRECTED RNA POLYM	1.65e+02	482	1	VAC3_HELPY	VACUOLATING CYTOTOXIN	1.65e+02
410	5	16.1	774	1	RRP3_INCB	RNA-DIRECTED RNA POLYM	1.65e+02	483	1	LT23_CAEEL	LET-23 RECEPTOR PROTEI	1.65e+02
411	5	16.1	777	1	FRZE_MYXAO	GLIDING MOTILITY REGUL	1.65e+02	484	1	DPOA_SCHPO	DNA POLYMERASE ALPHA C	1.65e+02
412	5	16.1	777	1	ISOA_FLASP	ISOAMYLASE PRECURSOR (	1.65e+02	485	1	EXPA_DROME	EXPANDED PROTEIN	1.65e+02
413	5	16.1	779	1	K6PF_RAT	6-PHOSPHOFUCTOKINASE,	1.65e+02	486	1	RPDQ_PVXHB	RNA REPLICATION PROTEI	1.65e+02
414	5	16.1	779	1	K6PF_RAT	6-PHOSPHOFUCTOKINASE,	1.65e+02	487	1	RPOD_OOSI	DNA REPLICATION RNA POLYM	1.65e+02
415	5	16.1	779	1	K6PF_HUMAN	6-PHOSPHOFUCTOKINASE,	1.65e+02	488	1	DPOA_DROME	DNA POLYMERASE ALPHA C	1.65e+02
416	5	16.1	780	1	STRI_RAT	STRIATIN	1.65e+02	489	1	MYSN_ACACA	MYOSIN II HEAVY CHAIN,	1.65e+02
417	5	16.1	781	1	K6PF_CANFA	6-PHOSPHOFUCTOKINASE,	1.65e+02	490	1	168K_RICTY	168 KD SURFACE-LAYER P	1.65e+02
418	5	16.1	784	1	K6PP_HUMAN	6-PHOSPHOFUCTOKINASE,	1.65e+02	491	1	FUR2_DROME	FURIN-LIKE PROTEASE 2	1.65e+02
419	5	16.1	791	1	TEX_BORPE	TEX PROTEIN	1.65e+02	492	1	POLN_FCVF9	NON-STRUCTURAL POLYPRO	1.65e+02
420	5	16.1	791	1	K6PP_RABIT	6-PHOSPHOFUCTOKINASE,	1.65e+02	493	1	DPOL_THEST	DNA POLYMERASE (EC 2.7	1.65e+02
421	5	16.1	800	1	ZFX_BOVIN	ZINC FINGER X-CHROMOSO	1.65e+02	494	1	MYSE_MESAO	MYOSIN HEAVY CHAIN, CA	1.65e+02
422	5	16.1	801	1	ZFY_HUMAN	ZINC FINGER Y-CHROMOSO	1.65e+02	495	1	MYSE_RAT	MYOSIN HEAVY CHAIN, CA	1.65e+02
423	5	16.1	802	1	NAB2_YEAST	NUCLEAR POLYADENYLATED	1.65e+02	496	1	RRPL_UUK	RNA POLYMERASE (EC 2.7	1.65e+02
424	5	16.1	802	1	PUR3_YEAST	PHOSPHORIBOSYLAMINE'-G	1.65e+02	497	1	RNA_DROME	RNA-DIRECTED RNA POLYM	1.65e+02
425	5	16.1	804	1	SVL_BACSU	LEUCYL-TRNA SYNTHETASE	1.65e+02	498	1	RRPL_HANTV	RNA-DIRECTED RNA POLYM	1.65e+02
426	5	16.1	820	1	FIBL_HUMAN	FIBRINOGEN-LIKE PROTEI	1.65e+02	499	1	RRPL_PUUMH	RNA-DIRECTED RNA POLYM	1.65e+02
427	5	16.1	826	1	PFUA_PSESP	FERRIC-PSEUDOBACTIN M1	1.65e+02	500	1	RRPB_TACV	RNA-DIRECTED RNA POLYM	1.65e+02
428	5	16.1	827	1	PTFL_RHOCA	MULTIPHOSPHORYL TRANSF	1.65e+02	501	1	RRPB_BEV	RNA-DIRECTED RNA POLYM	1.65e+02
429	5	16.1	827	1	SP2E_BACSU	STAGE II SPOREULATION P	1.65e+02	502	1	POLN_RHDV	NON-STRUCTURAL POLYPRO	1.65e+02
430	5	16.1	829	1	VIRA_AGR16	WIDE HOST RANGE VIRA P	1.65e+02	503	1	POLN_ONNVG	NON-STRUCTURAL POLYPRO	1.65e+02
431	5	16.1	829	1	VIRA_AGR16	WIDE HOST RANGE VIRA P	1.65e+02	504	1	NOTC_DROME	NEUROGENIC LOCUS NOTCH	1.65e+02
432	5	16.1	833	1	VIRA_AGR15	WIDE HOST RANGE VIRA P	1.65e+02	505	1	ZFH2_DROME	ZINC-FINGER PROTEIN 2	1.65e+02
433	5	16.1	838	1	SVY_MYCPN	VALYL-TRNA SYNTHETASE	1.65e+02	506	1	GENOME_POLYPROTEIN [CO	GENOME POLYPROTEIN [CO	1.65e+02
434	5	16.1	856	1	YW21_MYCTU	PROBABLE CATION-TRANSF	1.65e+02	507	1	LMA2_MOUSE	LAMININ ALPHA-2 CHAIN	1.65e+02
435	5	16.1	857	1	LOX3_SOYBN	SEED LIPOXYGENASE-3 (E	1.65e+02	508	1	ERY2_SACER	ERYTHRONOLIDE SYNTHASE	1.65e+02
436	5	16.1	861	1	GLGB_SOLTU	1,4-ALPHA-GLUCAN BRANC	1.65e+02	509	1	Y00A_BPT4	HYPOTHETICAL 8.1 KD PR	3.22e+03
437	5	16.1	863	1	YEJO_ECOLI	HYPOTHETICAL 91.2 KD P	1.65e+02	510	1	YDFZ_ECOLI	HYPOTHETICAL 7.3 KD PR	3.22e+03
438	5	16.1	877	1	SYA_SYNX3	ALANYL-TRNA SYNTHETASE	1.65e+02	511	1	YD28_METTH	HYPOTHETICAL TRANSSCRIP	3.22e+03
439	5	16.1	881	1	HELI_VZVD	PROBABLE HELICASE	1.65e+02	512	1	YGR8_CAEEL	HYPOTHETICAL 7.9 KD LI	3.22e+03
440	5	16.1	884	1	SYLM_SACDO	LEUCYL-TRNA SYNTHETASE	1.65e+02	513	1	YRR8_CAEEL	HYPOTHETICAL 6.9 KD PR	3.22e+03
441	5	16.1	895	1	ODPL_ALCEU	PYRUVATE DEHYDROGENASE	1.65e+02	514	1	Y055_NPVAC	HYPOTHETICAL 8.2 KD PR	3.22e+03
442	5	16.1	901	1	WGLB_GPCMV	GLYCOPROTEIN B PRECURS	1.65e+02	515	1	YFCB_SALTI	HYPOTHETICAL ADENINE-S	3.22e+03
443	5	16.1	907	1	ITH1_MOUSE	INTER-ALPHA-TRYPISIN IN	1.65e+02	516	1	YOH1_SERMA	HYPOTHETICAL 10.1 KD P	3.22e+03
444	5	16.1	915	1	PAC6_RAT	SERINE PROTEASE PC6 PR	1.65e+02	517	1	YXHE_BACSU	HYPOTHETICAL 11.9 KD P	3.22e+03
445	5	16.1	915	1	PAC6_MOUSE	SERINE PROTEASE PC6 PR	1.65e+02	518	1	YBHO_YEAST	VERY HYPOTHETICAL 11.6	3.22e+03
446	5	16.1	923	1	GYRA_PSEAE	DNA GYRASE SUBUNIT A (	1.65e+02	519	1	YATU_MYCGA	HYPOTHETICAL PROTEIN I	3.22e+03
447	5	16.1	936	1	PM3A_NICPL	PLASMA MEMBRANE ATPASE	1.65e+02	520	1	YATU_MYCGA	HYPOTHETICAL 12.4 KD P	3.22e+03
448	5	16.1	956	1	PM41_LYCES	PLASMA MEMBRANE ATPASE	1.65e+02	521	1	Y905_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
449	5	16.1	957	1	PM41_NICPL	PLASMA MEMBRANE ATPASE	1.65e+02	522	1	YFTA_ECOLI	12.7 KD PROTEIN IN SFH	3.22e+03
450	5	16.1	994	1	SYLM_NEUCR	LEUCYL-TRNA SYNTHETASE	1.65e+02	523	1	YGAM_ECOLI	HYPOTHETICAL 12.3 KD P	3.22e+03
451	5	16.1	996	1	VGNM_RCMV	GENOME POLYPROTEIN M (	1.65e+02	524	1	YRGL_LACLA	HYPOTHETICAL PROTEIN I	3.22e+03
452	5	16.1	996	1	PPOL_SARPE	POLY [ADP-RIBOSE] POLY	1.65e+02	525	1	YRGL_LACLA	HYPOTHETICAL PROTEIN I	3.22e+03
453	5	16.1	998	1	BGAL_LACIA	BETA-GALACTOSIDASE (EC	1.65e+02	526	1	YRGL_LACLA	HYPOTHETICAL 13.9 KD P	3.22e+03
454	5	16.1	999	1	DSG3_HUMAN	DESMOGLEIN 3 PRECURSOR	1.65e+02	527	1	YHIT_ACOBR	HYPOTHETICAL 13.2 KD H	3.22e+03
455	5	16.1	1005	1	Y321_MYCPN	HYPOTHETICAL LIPOPROTE	1.65e+02	528	1	YBNC_ECOLI	HYPOTHETICAL 14.2 KD P	3.22e+03
456	5	16.1	1013	1	PTPX_MACNE	PROTEIN-TYROSINE PHOSP	1.65e+02	529	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03
457	5	16.1	1015	1	PTPX_HUMAN	PROTEIN-TYROSINE PHOSP	1.65e+02	530	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03
458	5	16.1	1025	1	SLAP_CAUCR	S-LAYER PROTEIN (PARAC	1.65e+02	531	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03
459	5	16.1	1036	1	HPI_DEIRA	HEXAGONALLY PACKED INT	1.65e+02	532	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03
460	5	16.1	1043	1	MYSH_BOVIN	MYOSIN I HEAVY CHAIN-L	1.65e+02	533	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03
461	5	16.1	1049	1	RT31_ACTPL	RTX-III TOXIN DETERMIN	1.65e+02	534	1	VNS2_INBLA	NONSTRUCTURAL PROTEIN	3.22e+03

535	4	12.9	123	1	YJ54_YEAST	HYPOTHETICAL 13.6 KD P	3.22e+03	608	4	12.9	246	1	YJ1H_STRCG	HYPOTHETICAL 25.3 KD P	3.22e+03
536	4	12.9	125	1	Y007_BPL2	HYPOTHETICAL 14.0 KD P	3.22e+03	609	4	12.9	247	1	YBGI_ECOLI	HYPOTHETICAL 26.9 KD P	3.22e+03
537	4	12.9	128	1	YC35_ODOSI	HYPOTHETICAL 14.9 KD P	3.22e+03	610	4	12.9	247	1	Y335_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
538	4	12.9	129	1	YX31_CAEEL	HYPOTHETICAL 14.2 KD P	3.22e+03	611	4	12.9	250	1	YQEE_BACSU	PROBABLE N-ACETYLMURAM	3.22e+03
539	4	12.9	129	1	YGRS_STRCO	HYPOTHETICAL PROTEIN I	3.22e+03	612	4	12.9	251	1	YQ51_STRGC	HYPOTHETICAL ABC TRANS	3.22e+03
540	4	12.9	130	1	Y83B_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	613	4	12.9	251	1	YQ55_CAEEL	HYPOTHETICAL 28.6 KD P	3.22e+03
541	4	12.9	130	1	YAI1_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	614	4	12.9	251	1	YQ52_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
542	4	12.9	130	1	YHFU_ECOLI	HYPOTHETICAL 14.0 KD P	3.22e+03	615	4	12.9	252	1	YQIH_ECOLI	HYPOTHETICAL FIMBRIAL	3.22e+03
543	4	12.9	130	1	YGI0_YEAST	HYPOTHETICAL 15.0 KD P	3.22e+03	616	4	12.9	252	1	YKC6_CAEEL	HYPOTHETICAL 29.3 KD P	3.22e+03
544	4	12.9	131	1	YJG4_YEAST	HYPOTHETICAL 13.9 KD P	3.22e+03	617	4	12.9	255	1	YFUB_RHOVI	HYPOTHETICAL PROTEIN I	3.22e+03
545	4	12.9	132	1	Y747_CAEEL	HYPOTHETICAL 14.8 KD P	3.22e+03	618	4	12.9	255	1	Y05A_MYCTU	HYPOTHETICAL 27.3 KD P	3.22e+03
546	4	12.9	135	1	YF19_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03	619	4	12.9	256	1	YRSU_MYCLE	HYPOTHETICAL 28.1 KD P	3.22e+03
547	4	12.9	137	1	YEA_A_ECOLI	HYPOTHETICAL 15.5 KD P	3.22e+03	620	4	12.9	256	1	YK56_YEAST	HYPOTHETICAL 28.9 KD P	3.22e+03
548	4	12.9	139	1	YHT_BORBU	HYPOTHETICAL 15.9 KD H	3.22e+03	621	4	12.9	258	1	YK10_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
549	4	12.9	139	1	YPRE_HUMAN	IMMUNOGLOBULIN IOTA CH	3.22e+03	622	4	12.9	258	1	YK2F_SCHPO	HYPOTHETICAL 29.8 KD P	3.22e+03
550	4	12.9	140	1	Y4KN_RHISN	HYPOTHETICAL 15.7 KD P	3.22e+03	623	4	12.9	261	1	YK82_CAEEL	HYPOTHETICAL HISTONE P	3.22e+03
551	4	12.9	143	1	YORH_TTV1	HYPOTHETICAL 16.6 KD P	3.22e+03	624	4	12.9	261	1	Y558_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
552	4	12.9	143	1	YQ15_MYCTU	HYPOTHETICAL 15.6 KD P	3.22e+03	625	4	12.9	261	1	YR31_FRG3V	EARLY 31 KD PROTEIN	3.22e+03
553	4	12.9	147	1	YASK_MYCSM	HYPOTHETICAL 15.4 KD P	3.22e+03	626	4	12.9	262	1	YK23_METJA	HYPOTHETICAL ABC TRANS	3.22e+03
554	4	12.9	148	1	YCX3_ODOSI	HYPOTHETICAL 17.4 KD P	3.22e+03	627	4	12.9	263	1	YK23_HAEIN	HYPOTHETICAL ABC TRANS	3.22e+03
555	4	12.9	152	1	Y1ZG_BPT4	HYPOTHETICAL 17.1 KD P	3.22e+03	628	4	12.9	265	1	YQ01_MYCTU	HYPOTHETICAL 27.9 KD P	3.22e+03
556	4	12.9	156	1	YK04_PARTE	HYPOTHETICAL 18.8 KD P	3.22e+03	629	4	12.9	265	1	Y280_MYCGE	HYPOTHETICAL PROTEIN M	3.22e+03
557	4	12.9	162	1	YLF6_CAEEL	HYPOTHETICAL 18.5 KD P	3.22e+03	630	4	12.9	265	1	YIT6_YEAST	HYPOTHETICAL 28.9 KD P	3.22e+03
558	4	12.9	164	1	YQAD_ENTAG	HYPOTHETICAL 17.2 KD P	3.22e+03	631	4	12.9	266	1	YD37_SCHPO	VERY HYPOTHETICAL 31.0	3.22e+03
559	4	12.9	167	1	YRDA_BACSU	HYPOTHETICAL 18.9 KD P	3.22e+03	632	4	12.9	267	1	Y125_MYCCA	HYPOTHETICAL PROTEIN I	3.22e+03
560	4	12.9	168	1	YBPQ_BACSU	HYPOTHETICAL 19.3 KD P	3.22e+03	633	4	12.9	267	1	YATR_BACFI	HYPOTHETICAL ABC TRANS	3.22e+03
561	4	12.9	172	1	Y096_NPVOP	HYPOTHETICAL 19.4 KD P	3.22e+03	634	4	12.9	268	1	YOR3_LORDV	HYPOTHETICAL 28.5 KD P	3.22e+03
562	4	12.9	175	1	YFG_BBP2	PROBABLE TAIL FIBRE AS	3.22e+03	635	4	12.9	275	1	YCP1_BRAJA	PROBABLE SHORT-CHAIN T	3.22e+03
563	4	12.9	176	1	YQJB_BACSU	HYPOTHETICAL 19.7 KD P	3.22e+03	636	4	12.9	276	1	YPH1_METEX	HYPOTHETICAL PROTEIN I	3.22e+03
564	4	12.9	179	1	YEH6_YEAST	HYPOTHETICAL 20.7 KD P	3.22e+03	637	4	12.9	277	1	YQ23_ANTSP	HYPOTHETICAL 30.3 KD P	3.22e+03
565	4	12.9	180	1	YEB1_WA1ZE	ZEIN-BETA PRECURSOR (1	3.22e+03	638	4	12.9	277	1	Y233_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
566	4	12.9	180	1	YAF5_YEAST	HYPOTHETICAL 19.9 KD P	3.22e+03	639	4	12.9	278	1	Y4LA_RHISN	PUTATIVE SHORT-CHAIN T	3.22e+03
567	4	12.9	182	1	YCBQ_ECOLI	HYPOTHETICAL 19.9 KD P	3.22e+03	640	4	12.9	279	1	YKXA_LAMBD	REXA PROTEIN	3.22e+03
568	4	12.9	182	1	Y696_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	641	4	12.9	280	1	YHM7_YEAST	HYPOTHETICAL 33.1 KD P	3.22e+03
569	4	12.9	185	1	YALI_TRYBB	HYPOTHETICAL 22 KD PRO	3.22e+03	642	4	12.9	281	1	YH8G_AZACA	PROBABLE ABC TRANSPORT	3.22e+03
570	4	12.9	191	1	YK22_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03	643	4	12.9	281	1	Y373_MYCPN	HYPOTHETICAL PROTEIN M	3.22e+03
571	4	12.9	194	1	YGG8_YEAST	PUTATIVE MITOCHONDRIAL	3.22e+03	644	4	12.9	283	1	Y1JO_ECOLI	HYPOTHETICAL TRANSCRIP	3.22e+03
572	4	12.9	196	1	YG5C_YEAST	HYPOTHETICAL 22.3 KD P	3.22e+03	645	4	12.9	283	1	YK21_YEAST	HYPOTHETICAL 32.2 KD P	3.22e+03
573	4	12.9	197	1	YG77_YEAST	HYPOTHETICAL 22.8 KD P	3.22e+03	646	4	12.9	283	1	YK20_YEAST	HYPOTHETICAL 32.8 KD P	3.22e+03
574	4	12.9	200	1	YNU2_SHIFL	HYPOTHETICAL 22.8 KD P	3.22e+03	647	4	12.9	284	1	YKPN_CAEEL	PUTATIVE CUTICLE COLLA	3.22e+03
575	4	12.9	201	1	YV14_MYCTU	HYPOTHETICAL 20.9 KD P	3.22e+03	648	4	12.9	286	1	VNS4_RSVM	NONSTRUCTURAL PROTEIN	3.22e+03
576	4	12.9	201	1	YVPT_ADE04	ENDOPROTEASE (EC 3.4.2	3.22e+03	649	4	12.9	286	1	YGG8_ECOLI	HYPOTHETICAL 30.9 KD P	3.22e+03
577	4	12.9	205	1	Y1JF_ECOLI	HYPOTHETICAL 23.0 KD P	3.22e+03	650	4	12.9	287	1	Y280_MYCPN	HYPOTHETICAL PROTEIN M	3.22e+03
578	4	12.9	206	1	Y312_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	651	4	12.9	288	1	Y1GM_ECOLI	HYPOTHETICAL 32.4 KD P	3.22e+03
579	4	12.9	206	1	YG21_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03	652	4	12.9	288	1	Y143_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03
580	4	12.9	206	1	YVPT_ADE12	ENDOPROTEASE (EC 3.4.2	3.22e+03	653	4	12.9	290	1	Y873_METJA	HYPOTHETICAL ABC TRANS	3.22e+03
581	4	12.9	206	1	Y817_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	654	4	12.9	290	1	Y43A_MYCPN	HYPOTHETICAL LIPOPROTE	3.22e+03
582	4	12.9	209	1	YDGI_BACSU	PUTATIVE NAD(P)H NITRO	3.22e+03	655	4	12.9	291	1	YPE6_NPVLD	HYPOTHETICAL 33.9 KD P	3.22e+03
583	4	12.9	209	1	YVPT_ADE03	ENDOPROTEASE (EC 3.4.2	3.22e+03	656	4	12.9	292	1	YKX2_YEAST	HYPOTHETICAL 31.6 KD P	3.22e+03
584	4	12.9	210	1	YEO9_YEAST	HYPOTHETICAL 23.2 KD P	3.22e+03	657	4	12.9	292	1	YBQ_ECOLI	HYPOTHETICAL 30.8 KD P	3.22e+03
585	4	12.9	210	1	YQ53_CAEEL	HYPOTHETICAL 23.6 KD P	3.22e+03	658	4	12.9	294	1	YHBM_ECOLI	HYPOTHETICAL 33.6 KD P	3.22e+03
586	4	12.9	211	1	YAN8_SCHPO	HYPOTHETICAL 24.6 KD P	3.22e+03	659	4	12.9	295	1	YFLL_YEAST	HYPOTHETICAL 34.2 KD P	3.22e+03
587	4	12.9	217	1	YK46_YEAST	HYPOTHETICAL 24.6 KD P	3.22e+03	660	4	12.9	295	1	YKKB_BACSU	HYPOTHETICAL 32.9 KD P	3.22e+03
588	4	12.9	221	1	YK39_SCHPO	PUTATIVE ACETYLTRANSFER	3.22e+03	661	4	12.9	295	1	YRIF_ECOLI	HYPOTHETICAL 31.3 KD P	3.22e+03
589	4	12.9	225	1	Y810_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	662	4	12.9	300	1	YATE_RHISN	PROBABLE AMINO-ACID AB	3.22e+03
590	4	12.9	226	1	YHSH_HALMA	HYPOTHETICAL PROTEIN I	3.22e+03	663	4	12.9	306	1	YFDH_ECOLI	HYPOTHETICAL 34.6 KD P	3.22e+03
591	4	12.9	230	1	YF2A_STAAT	HYPOTHETICAL 26.9 KD P	3.22e+03	664	4	12.9	309	1	YFDB_BACSU	HYPOTHETICAL 34.0 KD P	3.22e+03
592	4	12.9	231	1	YDQJ_ECOLI	HYPOTHETICAL 24.5 KD P	3.22e+03	665	4	12.9	309	1	YQAB_SCHPO	HYPOTHETICAL 35.3 KD P	3.22e+03
593	4	12.9	231	1	YJCU_ECOLI	HYPOTHETICAL 26.1 KD P	3.22e+03	666	4	12.9	309	1	YAGE_ECOLI	HYPOTHETICAL 33.3 KD P	3.22e+03
594	4	12.9	232	1	YNF7_RHIME	HYPOTHETICAL 25.1 KD P	3.22e+03	667	4	12.9	309	1	Y1L1_CAEEL	HYPOTHETICAL 35.3 KD P	3.22e+03
595	4	12.9	232	1	YDFE_SCHPO	HYPOTHETICAL 26.1 KD P	3.22e+03	668	4	12.9	310	1	YAGD_ECOLI	HYPOTHETICAL 33.4 KD P	3.22e+03
596	4	12.9	234	1	YK02_CAEEL	HYPOTHETICAL 27.8 KD P	3.22e+03	669	4	12.9	311	1	YQBE_BACSU	HYPOTHETICAL 34.5 KD P	3.22e+03
597	4	12.9	234	1	YD57_SCHPO	HYPOTHETICAL 25.9 KD P	3.22e+03	670	4	12.9	311	1	YKDG_BACSU	PHAGE-LIKE ELEMENT PBS	3.22e+03
598	4	12.9	236	1	YIN1_STRAM	HYPOTHETICAL TRANSCRIP	3.22e+03	671	4	12.9	311	1	YD37_SCHPO	PHAGE-LIKE ELEMENT PBS	3.22e+03
599	4	12.9	236	1	YAA3_SCHPO	HYPOTHETICAL 26.9 KD P	3.22e+03	672	4	12.9	312	1	YVPE_NPVLD	POLYHEDRAL ENVELOPE PR	3.22e+03
600	4	12.9	237	1	YPA4_METTF	HYPOTHETICAL 26.5 KD P	3.22e+03	673	4	12.9	315	1	YQAM_YEAST	HYPOTHETICAL 34.9 KD P	3.22e+03
601	4	12.9	240	1	YH8G_ECOLI	PROBABLE ABC TRANSPORT	3.22e+03	674	4	12.9	317	1	YH91_YEAST	HYPOTHETICAL 35.9 KD P	3.22e+03
602	4	12.9	240	1	YFPH_ECOLI	HYPOTHETICAL 25.7 KD P	3.22e+03	675	4	12.9	319	1	YK52_SCHPO	HYPOTHETICAL 36.6 KD P	3.22e+03
603	4	12.9	241	1	YH8G_THIFE	PROBABLE ABC TRANSPORT	3.22e+03	676	4	12.9	320	1	YDJE_BACSU	HYPOTHETICAL SUGAR KIN	3.22e+03
604	4	12.9	243	1	YQBF_BACSU	HYPOTHETICAL 27.6 KD L	3.22e+03	677	4	12.9	321	1	YB1K_ECOLI	HYPOTHETICAL 33.4 KD P	3.22e+03
605	4	12.9	244	1	YKCD_CYAPA	PROBABLE ABC TRANSPORT	3.22e+03	678	4	12.9	321	1	YHDG_SALTY	HYPOTHETICAL 35.9 KD P	3.22e+03
606	4	12.9	244	1	YV24_MYCLE	HYPOTHETICAL 26.6 KD P	3.22e+03	679	4	12.9	322	1	YQ13_CAEEL	HYPOTHETICAL 35.9 KD P	3.22e+03
607	4	12.9	245	1	YIT8_YEAST	HYPOTHETICAL 28.4 KD P	3.22e+03	680	4	12.9	322	1	YK02_MARPO	HYPOTHETICAL 37.0 KD P	3.22e+03



827	4	12.9	529	1	Y119_NPVOP	HYPOTHETICAL 59.0 KD P	3.22e+03	900	1	VPS1_YEAST	VACUOLAR SORTING PROTE	3.22e+03
828	4	12.9	534	1	YE25_METH	O-SIALOGLYCOPROTEIN EN	3.22e+03	901	1	Y253_HUMAN	HYPOTHETICAL PROTEIN K	3.22e+03
829	4	12.9	535	1	YB30_METJA	HYPOTHETICAL PROTEIN M	3.22e+03	902	1	YHC4_YEAST	HYPOTHETICAL 80.1 KD P	3.22e+03
830	4	12.9	536	1	YEN1_SCHPO	HYPOTHETICAL 52.9 KD S	3.22e+03	903	1	YKA4_CAEEL	HYPOTHETICAL 82.8 KD P	3.22e+03
831	4	12.9	537	1	YW18_MYCTU	HYPOTHETICAL 59.9 KD P	3.22e+03	904	1	YMP8_CAEEL	HYPOTHETICAL 82.6 KD P	3.22e+03
832	4	12.9	541	1	YKC5_CAEEL	HYPOTHETICAL 55.6 KD P	3.22e+03	905	4	YCCS_ECOLI	HYPOTHETICAL 82.0 KD P	3.22e+03
833	4	12.9	545	1	YAO5_SCHPO	HYPOTHETICAL 61.1 KD P	3.22e+03	906	4	YMW6_YEAST	HYPOTHETICAL 82.0 KD P	3.22e+03
834	4	12.9	551	1	YIDE_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03	907	1	Y4NA_RHISN	PROBABLE PEPTIDASE Y4N	3.22e+03
835	4	12.9	554	1	YJJK_ECOLI	ABC TRANSPORTER ATP-BI	3.22e+03	908	4	YPR4_BACSU	HYPOTHETICAL HELICASE	3.22e+03
836	4	12.9	556	1	YJJK_HAEIN	ABC TRANSPORTER ATP-BI	3.22e+03	909	4	YBHJ_ECOLI	HYPOTHETICAL 81.5 KD P	3.22e+03
837	4	12.9	557	1	Y281_MYGE	HYPOTHETICAL PROTEIN M	3.22e+03	910	4	YAJ3_SCHPO	PUTATIVE ATP-DEPENDENT	3.22e+03
838	4	12.9	558	1	YK36_MYCTU	HYPOTHETICAL 53.7 KD P	3.22e+03	911	4	YBIL_ECOLI	PROBABLE TONB-DEPENDEN	3.22e+03
839	4	12.9	558	1	YK36_MYCTU	HYPOTHETICAL 64.9 KD P	3.22e+03	912	4	VPA_EPP2	REPLICATION GENE A PRO	3.22e+03
840	4	12.9	560	1	YICF_ECOLI	HYPOTHETICAL 63.2 KD P	3.22e+03	913	4	VP4_ROT89	OUTER CAPSID PROTEIN V	3.22e+03
841	4	12.9	563	1	YWO3_CAEEL	HYPOTHETICAL 63.6 KD P	3.22e+03	914	4	YQ04_MYCTU	HYPOTHETICAL 63.1 KD G	3.22e+03
842	4	12.9	569	1	YSY1_YEAST	HYPOTHETICAL 65.4 KD P	3.22e+03	915	4	YHJN_ECOLI	HYPOTHETICAL 86.0 KD P	3.22e+03
843	4	12.9	571	1	YXNC_PSEFL	ALPHA-L-ARABINOFURANOS	3.22e+03	916	4	YHCD_ECOLI	HYPOTHETICAL OUTER MEM	3.22e+03
844	4	12.9	575	1	YDGE_SCHPO	PUTATIVE FLAVOPROTEIN	3.22e+03	917	4	YEL4_YEAST	PUTATIVE 91.1 KD TRANS	3.22e+03
845	4	12.9	577	1	YIJP_ECOLI	HYPOTHETICAL 66.6 KD P	3.22e+03	918	4	YAFH_ECOLI	HYPOTHETICAL 87.0 KD P	3.22e+03
846	4	12.9	589	1	VP40_SCHVC	CAPSID PROTEIN P40 [CO	3.22e+03	919	4	Y1B2_MYCTU	PUTATIVE CATION-TRANS	3.22e+03
847	4	12.9	593	1	YHYA_BPH44	HYPOTHETICAL 65 KD PRO	3.22e+03	920	4	YMT9_YEAST	HYPOTHETICAL 91.7 KD T	3.22e+03
848	4	12.9	595	1	YB85_SCHPO	HYPOTHETICAL 65.3 KD T	3.22e+03	921	4	YBGO_ECOLI	HYPOTHETICAL OUTER MEM	3.22e+03
849	4	12.9	598	1	YF64_CAEEL	HYPOTHETICAL 67.0 KD P	3.22e+03	922	4	YOB3_NPVOP	HYPOTHETICAL 91.1 KD P	3.22e+03
850	4	12.9	601	1	WD66_PHYPO	66 KD STRESS PROTEIN (	3.22e+03	923	4	YC81_CHLVU	HYPOTHETICAL 94.7 KD P	3.22e+03
851	4	12.9	602	1	Y213_METJA	HYPOTHETICAL MCM-TYPE	3.22e+03	924	4	VP41_CAEEL	VACUOLAR ASSEMBLY PROT	3.22e+03
852	4	12.9	603	1	YQ3A_MYCTU	HYPOTHETICAL GLYCINE-R	3.22e+03	925	4	YRR6_CAEEL	HYPOTHETICAL 95.2 KD P	3.22e+03
853	4	12.9	604	1	YEJA_ECOLI	HYPOTHETICAL 69.7 KD P	3.22e+03	926	4	VPH1_YEAST	VACUOLAR ATP SYNTHASE	3.22e+03
854	4	12.9	604	1	Y316_YEAST	HYPOTHETICAL 68.4 KD P	3.22e+03	927	4	VP2_ROTPC	RNA-BINDING PROTEIN VP	3.22e+03
855	4	12.9	607	1	YA27_SCHPO	HYPOTHETICAL 68.8 KD P	3.22e+03	928	4	YQ66_NPVOP	HYPOTHETICAL 98.6 KD P	3.22e+03
856	4	12.9	608	1	Y1P5_CAEEL	HYPOTHETICAL 66.5 KD P	3.22e+03	929	4	YBGG_ECOLI	HYPOTHETICAL 100.0 KD	3.22e+03
857	4	12.9	608	1	YD56_YEAST	PUTATIVE MULTICOPPER O	3.22e+03	930	4	Y277_MYCPN	HYPOTHETICAL PROTEIN M	3.22e+03
858	4	12.9	610	1	YEM6_YEAST	PROBABLE ATP-DEPENDENT	3.22e+03	931	4	YH1H_ECOLI	HYPOTHETICAL ABC TRANS	3.22e+03
859	4	12.9	611	1	YD3M_HERAU	HYPOTHETICAL 68.4 KD P	3.22e+03	932	4	VP3_EHDV1	VP3 CORE PROTEIN.	3.22e+03
860	4	12.9	615	1	YCHO_YEAST	HYPOTHETICAL 68.9 KD P	3.22e+03	933	4	VP3_EHDVA	VP3 CORE PROTEIN.	3.22e+03
861	4	12.9	622	1	YFEL_YEAST	PUTATIVE MULTICOPPER O	3.22e+03	934	4	Y188_HUMAN	HYPOTHETICAL PROTEIN K	3.22e+03
862	4	12.9	623	1	YAK8_SCHPO	PUTATIVE MULTICOPPER O	3.22e+03	935	4	YB75_YEAST	HYPOTHETICAL 101.2 KD	3.22e+03
863	4	12.9	623	1	YF22_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03	936	4	VP3_BTIV17	VP3 CORE PROTEIN (MAJO	3.22e+03
864	4	12.9	626	1	YQ6J_MYCTU	HYPOTHETICAL 67.9 KD P	3.22e+03	937	4	VP3_BTIV5	VP3 CORE PROTEIN (MAJO	3.22e+03
865	4	12.9	629	1	YD4B_SCHPO	HYPOTHETICAL 71.3 KD P	3.22e+03	938	4	VP3_BTIV11	VP3 CORE PROTEIN (MAJO	3.22e+03
866	4	12.9	633	1	YR45_CAEEL	HYPOTHETICAL 72.9 KD P	3.22e+03	939	4	VP3_BTIV1A	VP3 CORE PROTEIN (MAJO	3.22e+03
867	4	12.9	633	1	Y1J2_YEAST	HYPOTHETICAL 71.0 KD P	3.22e+03	940	4	VP3_BTIV2A	VP3 CORE PROTEIN (MAJO	3.22e+03
868	4	12.9	633	1	YN33_YEAST	HYPOTHETICAL 73.0 KD P	3.22e+03	941	4	YCL7_HAEIN	PROBABLE TONB-DEPENDEN	3.22e+03
869	4	12.9	637	1	YAO7_SCHPO	PUTATIVE 73.4 KD TRANS	3.22e+03	942	4	YK62_CAEEL	HYPOTHETICAL 105.4 KD	3.22e+03
870	4	12.9	638	1	YHES_HAEIN	HYPOTHETICAL ABC TRANS	3.22e+03	943	4	YNE6_CAEEL	PROBABLE 3',5'-CYCLIC	3.22e+03
871	4	12.9	639	1	YBE2_YEAST	PUTATIVE TRANSPORTER Y	3.22e+03	944	4	YNAQ_CAEEL	HYPOTHETICAL 105.5 KD	3.22e+03
872	4	12.9	642	1	YBBD_BACSU	HYPOTHETICAL 70.6 KD P	3.22e+03	945	4	Y321_MYCGE	HYPOTHETICAL LIPOPROTE	3.22e+03
873	4	12.9	642	1	YG22_YEAST	HYPOTHETICAL 71.3 KD P	3.22e+03	946	4	YNM7_YEAST	HYPOTHETICAL 109.8 KD	3.22e+03
874	4	12.9	644	1	YQ44_CAEEL	PROBABLE G PROTEIN-COU	3.22e+03	947	4	VP2_BTIV7	OUTER CAPSID PROTEIN V	3.22e+03
875	4	12.9	644	1	YED0_YEAST	HEAT SHOCK PROTEIN 70	3.22e+03	948	4	VP2_BTIV11	OUTER CAPSID PROTEIN V	3.22e+03
876	4	12.9	644	1	VP4_BTIV2A	VP4 CORE PROTEIN.	3.22e+03	949	4	YBS3_YEAST	HYPOTHETICAL 108.0 KD	3.22e+03
877	4	12.9	647	1	YIK1_YEAST	HYPOTHETICAL 72.7 KD P	3.22e+03	950	4	YSX7_CAEEL	HYPOTHETICAL 113.1 KD	3.22e+03
878	4	12.9	648	1	YQ95_TREPA	HYPOTHETICAL PROTEIN T	3.22e+03	951	4	YD30_YEAST	HYPOTHETICAL 112.2 KD	3.22e+03
879	4	12.9	651	1	YHJK_ECOLI	HYPOTHETICAL 73.1 KD P	3.22e+03	952	4	YAA2_SCHPO	HYPOTHETICAL 111.5 KD	3.22e+03
880	4	12.9	653	1	YDCP_ECOLI	PUTATIVE PROTEASE IN T	3.22e+03	953	4	YGZ1_YEAST	HYPOTHETICAL 113.9 KD	3.22e+03
881	4	12.9	656	1	YAN9_SCHPO	HYPOTHETICAL 73.1 KD P	3.22e+03	954	4	VP3_RDV	MAJOR 114 KD STRUCTURA	3.22e+03
882	4	12.9	657	1	YDGL_SCHPO	HYPOTHETICAL PROTEIN C	3.22e+03	955	4	YRD3_CAEEL	HYPOTHETICAL 112.3 KD	3.22e+03
883	4	12.9	661	1	YN84_YEAST	HYPOTHETICAL 75.4 KD P	3.22e+03	956	4	YQ75_MYCGE	HYPOTHETICAL PROTEIN M	3.22e+03
884	4	12.9	666	1	YEA7_YEAST	HYPOTHETICAL 72.5 KD P	3.22e+03	957	4	YQ75_MYCPN	HYPOTHETICAL PROTEIN H	3.22e+03
885	4	12.9	668	1	YKCA_CAEEL	HYPOTHETICAL 75.0 KD P	3.22e+03	958	4	YH95_HAEIN	HYPOTHETICAL PROTEIN H	3.22e+03
886	4	12.9	683	1	YK87_YEAST	HYPOTHETICAL 78.3 KD P	3.22e+03	959	4	YD55_SCHPO	PROBABLE CATION-TRANS	3.22e+03
887	4	12.9	684	1	YV18_MYCTU	HYPOTHETICAL 57.3 KD P	3.22e+03	960	4	YDOF_SCHPO	HYPOTHETICAL PROTEIN C	3.22e+03
888	4	12.9	686	1	YIK5_YEAST	HYPOTHETICAL 78.0 KD P	3.22e+03	961	4	YG35_YEAST	HYPOTHETICAL 117.0 KD	3.22e+03
889	4	12.9	688	1	YJ80_YEAST	HYPOTHETICAL 80.2 KD P	3.22e+03	962	4	YHIV_ECOLI	HYPOTHETICAL 111.5 KD	3.22e+03
890	4	12.9	690	1	YQ23_NPVAC	HYPOTHETICAL 79.9 KD P	3.22e+03	963	4	YR71_CAEEL	HYPOTHETICAL 118.2 KD	3.22e+03
891	4	12.9	692	1	YNUC_MABVM	NUCLEOPROTEIN (NUCLEOC	3.22e+03	964	4	YAF3_SCHPO	HYPOTHETICAL 118.6 KD	3.22e+03
892	4	12.9	692	1	VP3_ROTPC	INNER CORE PROTEIN VP3	3.22e+03	965	4	YBDE_ECOLI	HYPOTHETICAL 114.7 KD	3.22e+03
893	4	12.9	694	1	Y4BA_RHISN	HYPOTHETICAL 78.7 KD P	3.22e+03	966	4	YHV4_YEAST	HYPOTHETICAL 123.0 KD	3.22e+03
894	4	12.9	694	1	YB9C_YEAST	HYPOTHETICAL 79.0 KD P	3.22e+03	967	4	YXNY_CLOTM	ENDO-1,4-BETA-XYLANASE	3.22e+03
895	4	12.9	695	1	YNUC_MABVP	NUCLEOPROTEIN (NUCLEOC	3.22e+03	968	4	YIL2_YEAST	HYPOTHETICAL 123.6 KD	3.22e+03
896	4	12.9	698	1	YJBH_ECOLI	HYPOTHETICAL 78.5 KD L	3.22e+03	969	4	YE20_METJA	HYPOTHETICAL PROTEIN M	3.22e+03
897	4	12.9	699	1	YQOA_CAEEL	HYPOTHETICAL 76.5 KD P	3.22e+03	970	4	YJEP_ECOLI	HYPOTHETICAL 123.8 KD	3.22e+03
898	4	12.9	700	1	YDHV_ECOLI	HYPOTHETICAL 77.9 KD P	3.22e+03	971	4	YKDB_YEAST	PUTATIVE 128.2 KD TRAN	3.22e+03
899	4	12.9	704	1	YQZ9_YEAST	HYPOTHETICAL 82.8 KD P	3.22e+03	972	1	YXN5_CAEEL	HYPOTHETICAL 139.4 KD	3.22e+03

973 4 12.9 1226 1 YCS3\_YEAST HYPOTHETICAL 138.5 KD 3.22e+03  
974 4 12.9 1238 1 YN13\_YEAST HYPOTHETICAL 141.1 KD 3.22e+03  
975 4 12.9 1260 1 YAOE\_SCHPO HYPOTHETICAL 138.8 KD 3.22e+03  
976 4 12.9 1327 1 YW11\_MCTU HYPOTHETICAL 145.8 KD 3.22e+03  
977 4 12.9 1341 1 YL78\_YEAST PUTATIVE 151.3 KD TRAN 3.22e+03  
978 4 12.9 1345 1 YH00\_YEAST HYPOTHETICAL 149.7 KD 3.22e+03  
979 4 12.9 1381 1 YBET\_YEAST HYPOTHETICAL 150.8 KD 3.22e+03  
980 4 12.9 1411 1 YM42\_YEAST HYPOTHETICAL 162.7 KD 3.22e+03  
981 4 12.9 1411 1 YK63\_CAEEL HYPOTHETICAL 157.6 KD 3.22e+03  
982 4 12.9 1478 1 YAWB\_SCHPO PROBABLE ATP-DEPENDENT 3.22e+03  
983 4 12.9 1524 1 Y133\_HUMAN HYPOTHETICAL PROTEIN K 3.22e+03  
984 4 12.9 1540 1 YDBA\_ECOLI HYPOTHETICAL PROTEIN N 3.22e+03  
985 4 12.9 1559 1 YL15\_YEAST PROBABLE ATP-DEPENDENT 3.22e+03  
986 4 12.9 1722 1 YP93\_CAEEL HYPOTHETICAL HELICASE 3.22e+03  
987 4 12.9 1733 1 YNUA\_PVKA PROBABLE NUCLEAR ANTIG 3.22e+03  
988 4 12.9 1750 1 Y832\_METJA HYPOTHETICAL PROTEIN M 3.22e+03  
989 4 12.9 1755 1 YJZ7\_YEAST TRANSPOSON TY1 PROTEIN 3.22e+03  
990 4 12.9 1876 1 Y06K\_MCTU PROBABLE POLYKETIDE SY 3.22e+03  
991 4 12.9 1895 1 YLK3\_CAEEL PUTATIVE SERINE/THREON 3.22e+03  
992 4 12.9 2052 1 YDOB\_SCHPO HYPOTHETICAL 234.0 KD 3.22e+03  
993 4 12.9 2054 1 YCF2\_PINTH HYPOTHETICAL 244.6 KD 3.22e+03  
994 4 12.9 2131 1 YCF2\_SPIOL HYPOTHETICAL 250 KD PR 3.22e+03  
995 4 12.9 2216 1 YCF2\_EFIV1 HYPOTHETICAL 260 KD PR 3.22e+03  
996 4 12.9 2280 1 YCF2\_TOBAC HYPOTHETICAL 267 KD PR 3.22e+03  
997 4 12.9 2329 1 YS89\_CAEEL HYPOTHETICAL 254.3 KD 3.22e+03  
998 4 12.9 2334 1 YWAP\_BACSU WALL-ASSOCIATED PROTEI 3.22e+03  
999 4 12.9 3051 1 YNX3\_CAEEL HYPOTHETICAL PROTEIN T 3.22e+03  
1000 4 12.9 3343 1 YOG7\_CAEEL HYPOTHETICAL 375.7 KD 3.22e+03

## ALIGNMENTS

RESULT 1 STANDARD; PRT: 353 AA.  
ID YW17\_BORHE  
AC P32777;  
DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE VARIABLE MAJOR OUTER MEMBRANE LIPOPROTEIN 17 PRECURSOR.  
GN YW17.  
OS BORRELIA HERMSII.  
OG PLASMID.  
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SSP\_HSL SEROTYPE 17;  
RX MEDLINE: 9313110.  
RA RESTREPO B.I., KITTE T., CARTER C.J., INFANTE D., BARBOUR A.G.;  
RT "Subtelomeric expression regions of Borrelia hermsii linear plasmids  
are highly polymorphic."  
RL MOL. MICROBIOL. 6:3299-3311(1992).  
CC -!- FUNCTION: SERVES TO AVOID THE HOST IMMUNE RESPONSE BY CHANGING  
FROM ONE SURFACE EXPOSED VMP TO ANOTHER.  
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
ANCHOR.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L04788; G144024;  
DR PROSITE; P500013; PROKAR\_LIPOPROTEIN; 1.  
DR PFAM; PF00921; Lipoprotein\_2; 1.  
KW OUTER MEMBRANE; LIPOPROTEIN; SIGNAL; PLASMID.  
FT SIGNAL 1 18 PROBABLE  
FT CHAIN 19 353 VARIABLE MAJOR OUTER MEMBRANE  
FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).  
FT SEQUENCE 353 AA: 36460 MW; A0ACAF35 CRC32;  
SQ

Query Match 22.6%; Score 7; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 8.25e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 170 ADATKDD 176

QY 6 ADATKDD 12

## RESULT 2

ID ECIC\_BOVIN STANDARD; PRT: 903 AA.  
AC P54281;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE EPIHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE  
CHANNEL).  
DE BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TRACHEA;  
RX MEDLINE: 96125078.  
RA CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,  
ARRATE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;  
RT "Cloning of an epithelial chloride channel from bovine trachea."  
RL J. BIOL. CHEM. 270:31016-31026(1995).  
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
TRANSEPIHELIAL TRANSPORT.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- TISSUE SPECIFICITY: TRACHEA.  
CC -!- PTM: PHOSPHORYLATED BY CAM-KINASE II.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U36445; G1184066;  
DR IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;  
KW CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.  
FT TRANSMEM 7 27  
FT TRANSMEM 331 351  
FT TRANSMEM 617 637  
FT TRANSMEM 883 903  
FT CARBOHYD 75 75  
FT CARBOHYD 278 278  
FT CARBOHYD 360 360  
FT CARBOHYD 372 372  
FT CARBOHYD 504 504  
FT CARBOHYD 515 515  
FT CARBOHYD 688 688  
FT CARBOHYD 811 811  
FT CARBOHYD 816 816  
FT CARBOHYD 842 842  
FT CARBOHYD 857 857  
SQ SEQUENCE 903 AA; 100305 MW; 65E6EC1C CRC32;

Query Match 22.6%; Score 7; DB 1; Length 903;  
Best Local Similarity 100.0%; Pred. No. 8.25e-02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 651 DNGACAD 657

QY 1 DNGACAD 7

RESULT 3  
 ID CYMS\_RAT STANDARD; PRT; 92 AA.  
 AC P04166;  
 DT 20-MAR-1987 (REL. 04, CREATED)  
 DT 20-MAR-1987 (REL. 04, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE CYTOCHROME B5, OUTER MITOCHONDRIAL MEMBRANE (FRAGMENT).  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE; 83182449.  
 RA LEDERER F., GHIR R., GUIARD B., CORTIAL S., ITO A.;  
 RT "Two homologous cytochromes b5 in a single cell.";  
 RL EUR. J. BIOCHEM. 132:95-102(1983).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RX MEDLINE; 97128656.  
 RA RODRIGUEZ-WARANON M.J., QIU F., STARK R.E., WHITE S.P., ZHANG X.,  
 RA FOUNDLING S.I., RODRIGUEZ V., SCHILLING C.L. III, BUNCE R.A.,  
 RA RIVERA M.;  
 RT "13C NMR spectroscopic and X-ray crystallographic study of the role  
 played by mitochondrial cytochrome b5 heme propionates in the  
 electrostatic binding to cytochrome c.";  
 RL BIOCHEMISTRY 35:16378-16390(1996).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE; 96145212.  
 RA RIVERA M., SEETHARAMAN R., GIRDHAR D., WIRTZ M., ZHANG X., WANG X.,  
 RA WHITE S.;  
 RT "The reduction potential of cytochrome b5 is modulated by its exposed  
 heme edge.";  
 RL BIOCHEMISTRY 37:1485-1494(1998).  
 CC -!- FUNCTION: CYTOCHROME B5 IS A MEMBRANE BOUND HEMOPROTEIN WHICH  
 CC FUNCTION AS AN ELECTRON CARRIER FOR SEVERAL MEMBRANE BOUND  
 CC OXYGENASES.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.  
 CC -!- THIS OUTER MITOCHONDRIAL MEMBRANE CYTOCHROME B5, IS SIMILAR TO,  
 CC BUT NOT IDENTICAL WITH MICROSMAL CYTOCHROME B5.  
 DR PIR; A00173; CBR5M.  
 DR PDB; 1BSM; 12-MAR-97.  
 DR PDB; 1AWP; 18-NOV-98.  
 DR PROSITE; PS00191; CYTOCHROME\_B5; 1.  
 DR PFAM; PF00173; heme1; 1.  
 DR ELECTRON TRANSPORT; MITOCHONDRON; TRANSMEMBRANE; HEME; 3D-STRUCTURE.  
 FT DOMAIN 1 92 HEME-BINDING.  
 FT BINDING 44 44 HEME LIGAND (BY SIMILARITY).  
 FT BINDING 68 68 HEME LIGAND (BY SIMILARITY).  
 FT NON\_TER 92 92  
 SQ SEQUENCE 92 AA; 10378 MW; 2F7C6558 CRC32;  
 Query Match 19.4%; Score 6; DB 1; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 55 AGADAT 60  
 QY [1]  
 ID ATND\_RAT STANDARD; PRT; 279 AA.  
 AC Q63377;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-3 CHAIN (EC 3.6.1.37)  
 DE (SODIUM/POTASSIUM-DEPENDENT ATPASE) (ATPB-3).  
 GN ATP1B3.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,  
 RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,  
 RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,  
 RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKIY M.,  
 RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;  
 OS METHANOCOCCUS JANNASCHII.  
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;  
 OC METHANOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE; 96337999.  
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,  
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,  
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,  
 RA OVERBECK R., KIRKNESS E.F.,



"Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";  
SCIENCE 273:1058-1073(1996).  
CC -!- CATALYTIC ACTIVITY: L1-2,6-DIAMINOHEPTANEDIOATE = MESO-DIAMINOHEPTANEDIOATE.  
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF LYSINE FROM ASPARTATE SEMIALDEHYDE.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE DIAMINOPIMELATE EPIMERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U67554; G1591758; -.  
DR TIGR: M1119; -.  
DR PROSITE: PS01326; DAP\_EPIMERASE; 1.  
KW ISOMERASE; LYSINE BIOSYNTHESIS.  
FT ACT-SITE 76 76 BY SIMILARITY.  
FT ACT-SITE 237 237 BY SIMILARITY.  
SQ SEQUENCE 295 AA; 33442 MW; 64AF6902 CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 295;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 272 KDG VY 277  
| | | | |  
QY 10 KDG VY 15  
  
RESULT 6  
ID GSPK-ECOLI STANDARD; PRT; 327 AA.  
AC P45762;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE PROBABLE GENERAL SECRETION PATHWAY PROTEIN K.  
GN GSPK.  
OS ESCHERICHIA COLI.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE; 97426617.  
RA BLATTNER F.R.; PLUNKETT G. III; BLOCH C.A.; PERNA N.T.; BURLAND V.,  
RA RILEY M.; COLLADO-VIDES J.; GLASNER F.D.; ROSE C.K.; MAYHEW G.F.,  
RA GREGOR J.; DAVIS N.W.; KIRKPATRICK H.A.; GOEDEN M.A.; ROSE D.J.,  
RA MAU B.; SHAO Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL SCIENCE 277:1453-1474(1997).  
CC -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE EXPORT OF PROTEINS (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE EXEK/PULK/OUTK/ACPX FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U18997; G606266; -.  
DR EMBL: AF000409; G1789729; -.  
DR ECOGENE; EGI2894; GSPK.  
KW TRANSPORT; TRANSMEMBRANE; INNER MEMBRANE.

FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 7 27 POTENTIAL.  
FT DOMAIN 28 327 PERIPLASMIC (POTENTIAL).  
SQ SEQUENCE 327 AA; 37646 MW; 7EFC407 CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 327;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 285 YSR YFT 290  
| | | | |  
QY 15 YSR YFT 20  
  
RESULT 7  
ID YQ51-CAEEL STANDARD; PRT; 328 AA.  
AC Q09461;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PUTATIVE MITOCHONDRIAL CARRIER C16C10.1.  
GN C16C10.1.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNITEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA LLOYD C.;  
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.  
CC -!- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z46787; E1344073; -.  
DR WORMPEP; C16C10.1; CE01489.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
DR PFAM: PF00153; mito\_carr; 2.  
KW HYPOTHETICAL PROTEIN; MITOCHONDRION; REPEAT; TRANSMEMBRANE;  
KW TRANSPORT.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 146 166 POTENTIAL.  
FT TRANSMEM 237 257 POTENTIAL.  
SQ SEQUENCE 328 AA; 36743 MW; 60D6B092 CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 328;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 111 YFT YTD 116  
| | | | |  
QY 18 YFT YTD 23  
  
RESULT 8  
ID SAOX-ECOLI STANDARD; PRT; 372 AA.  
AC P40874; Q47144;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE PUTATIVE SARCOSINE OXIDASE (EC 1.5.3.1).  
GN SOLA.  
OS ESCHERICHIA COLI.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
RN [1]  
RP SEQUENCE FROM N.A.



CC STRAIN-K12 / W3110;  
RA KOYAMA Y., OHMORI H.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
CC SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE; 97426617.  
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
RA MAU B., SHAO Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL SCIENCE 277:1453-1474(1997).  
CC -!- CATALYTIC ACTIVITY: SARCOSINE + H(2)O + O(2) = GLYCINE +  
CC FORMALDEHYDE + H(2)O(2).  
CC -!- COFACTOR: FAD FLAVOPROTEIN (1 MOLE OF FAD PER MOLE OF ENZYME).  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- SIMILARITY: STRONG. TO OTHER SARCOSINE OXIDASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D31709; G1236737; -.  
DR EMBL; AF000207; G1787298; -.  
DR ECGENE; EG12669; SOLA.  
KW OXIDOREDUCTASE; FAD; FLAVOPROTEIN.  
FT NP\_BIND 4 34 FAD (ADP PART) (POTENTIAL).  
SQ SEQUENCE 372 AA; 40902 MW; 7B29793F CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 372;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 227 GRYSVK 232  
QY 26 GRYSVK 31  
  
RESULT 9  
ID CARA\_PSEAE STANDARD; PRT; 378 AA.  
AC P38098;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE CARBAMOYL-PHOSPHATE SYNTHASE SMALL CHAIN (EC 6.3.5.5) (CARBAMOYL-  
DE PHOSPHATE SYNTHETASE GLUTAMINE CHAIN).  
GN CARA.  
OS PSEUDOMONAS AERUGINOSA.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;  
OC PSEUDOMONAS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-ATCC 15692 / PA01;  
RX MEDLINE; 94222830.  
RA KWON D.H., LU C.D., WALTHALL D.A., BROWN T.M., HOUGHTON J.E.,  
RA ABDELAL A.T.;  
RT "Structure and regulation of the carab operon in Pseudomonas  
RT aeruginosa and Pseudomonas stutzeri: no untranslated region exists.";  
RL J. BACTERIOL. 176:2532-2542(1994).  
RN [2]  
RP PRELIMINARY SEQUENCE FROM N.A.  
RX MEDLINE; 90130295.  
RA WONG S.C., ABDELAL A.T.;  
RT "Unorthodox expression of an enzyme: evidence for an untranslated  
RT region within carA from Pseudomonas aeruginosa.";  
RL J. BACTERIOL. 172:630-642(1990).  
CC -!- CATALYTIC ACTIVITY: 2 ATP + GLUTAMINE + CO(2) + H(2)O = 2 ADP +  
CC ORTHOPHOSPHATE + GLUTAMATE + CARBAMOYL PHOSPHATE.  
CC -!- PATHWAY: INVOLVED IN BOTH ARGININE AND PYRIMIDINE BIOSYNTHESIS.  
CC -!- SUBUNIT: COMPOSED OF TWO CHAINS; THE SMALL (OR GLUTAMINE) CHAIN  
CC PROMOTES THE HYDROLYSIS OF GLUTAMINE TO AMMONIA, WHICH IS USED  
CC BY THE LARGE (OR AMMONIA) CHAIN TO SYNTHESIZE CARBAMOYL PHOSPHATE.  
CC -!- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-1 GLUTAMINE  
CC AMIDOTRANSFERASES.  
CC -!- CAUTION: REF.2 SEQUENCE WAS ORIGINALLY THOUGHT TO ORIGINATE FROM  
CC PSEUDOMONAS AERUGINOSA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U04992; G451651; -.  
DR EMBL; U81259; G1750385; -.  
DR PROSITE; PS00442; GATASE\_TYPE\_I; 1.  
DR PFAM; PF00117; GATase; 1.  
DR PFAM; PF00988; CPSase\_sm\_chain; 1.  
DR HSP; P00907; LJDB.  
KW ARGININE BIOSYNTHESIS; PYRIMIDINE BIOSYNTHESIS; LIGASE.  
FT DOMAIN 1 189 CPSASE.  
FT DOMAIN 190 378 GLUTAMINE AMIDOTRANSFERASE.  
FT ACT\_SITE 269 269 GATASE (BY SIMILARITY).  
SQ SEQUENCE 378 AA; 40828 MW; 8C4A6DFD CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 378;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 134 AGADAT 139  
QY 4 AGADAT 9  
  
RESULT 10  
ID CARA\_PSEST STANDARD; PRT; 384 AA.  
AC P38099; P20597;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE CARBAMOYL-PHOSPHATE SYNTHASE SMALL CHAIN (EC 6.3.5.5) (CARBAMOYL-  
DE PHOSPHATE SYNTHETASE GLUTAMINE CHAIN).  
GN CARA.  
OS PSEUDOMONAS STUTZERI (PSEUDOMONAS PERFECTOMARINA).  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;  
OC PSEUDOMONAS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-JM300;  
RX MEDLINE; 94222830.  
RA KWON D.H., LU C.D., WALTHALL D.A., BROWN T.M., HOUGHTON J.E.,  
RA ABDELAL A.T.;  
RT "Structure and regulation of the carab operon in Pseudomonas  
RT aeruginosa and Pseudomonas stutzeri: no untranslated region exists.";  
RL J. BACTERIOL. 176:2532-2542(1994).  
RN [2]  
RP PRELIMINARY SEQUENCE FROM N.A.  
RX MEDLINE; 90130295.  
RA WONG S.C., ABDELAL A.T.;  
RT "Unorthodox expression of an enzyme: evidence for an untranslated  
RT region within carA from Pseudomonas aeruginosa.";  
RL J. BACTERIOL. 172:630-642(1990).  
CC -!- CATALYTIC ACTIVITY: 2 ATP + GLUTAMINE + CO(2) + H(2)O = 2 ADP +  
CC ORTHOPHOSPHATE + GLUTAMATE + CARBAMOYL PHOSPHATE.  
CC -!- PATHWAY: INVOLVED IN BOTH ARGININE AND PYRIMIDINE BIOSYNTHESIS.  
CC -!- SUBUNIT: COMPOSED OF TWO CHAINS; THE SMALL (OR GLUTAMINE) CHAIN  
CC PROMOTES THE HYDROLYSIS OF GLUTAMINE TO AMMONIA, WHICH IS USED  
CC BY THE LARGE (OR AMMONIA) CHAIN TO SYNTHESIZE CARBAMOYL PHOSPHATE.  
CC -!- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-1 GLUTAMINE  
CC AMIDOTRANSFERASES.  
CC -!- CAUTION: REF.2 SEQUENCE WAS ORIGINALLY THOUGHT TO ORIGINATE FROM  
CC PSEUDOMONAS AERUGINOSA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL: U04993; G451655; -  
DR EMBL: M33818; G551926; -  
CC PIR: A35111; A35111.  
DR PROSITE: PS00442; GATASE\_TYPE\_I; 1.  
DR PFAM: PF00117; GATase; 1.  
DR PFAM: PF00988; CPSase\_sm\_chain; 1.  
DR HSP: P00907; 1JDB  
KW ARGinine BIOSYNTHESIS; PYRIMIDINE BIOSYNTHESIS; LIQASE.  
FT DOMAIN 1 193 CPSASE.  
FT ACT\_SITE 194 384  
FT ACT\_SITE 273 273  
FT CONFLICT 33 33  
FT CONFLICT 73 73  
FT CONFLICT 75 75  
FT CONFLICT 104 104  
FT CONFLICT 234 234  
FT CONFLICT 279 279  
FT CONFLICT 324 325  
FT CONFLICT 353 353  
SQ SEQUENCE 384 AA: 41488 MW: 12724AA4 CRC32;  
Query Match 19.4%; Score 6; DB 1; Length 384;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 138 AGADAT 143  
QY 4 AGADAT 9  
RESULT 11  
ID COAT\_TBSVB STANDARD; PRT: 387 AA.  
AC P11795;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DE COAT PROTEIN.  
OS TOMATO BUSHY STUNT VIRUS (STRAIN BS-3) (TBSV).  
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; TOMBUSVIRIDAE;  
OC TOMBUSVIRUS.  
RN [1]  
RP SEQUENCE.  
RX MEDLINE: 85009723.  
RA HOPPER P., HARRISON S.C., SAUER R.T.;  
RT "Structure of tomato bushy stunt virus. V. Coat protein sequence  
determination and its structural implications."  
RL J. MOL. BIOL. 177:701-713(1984).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE: 84064870.  
RA OLSON A.J., BRICOGNE G., HARRISON S.C.;  
RT "Structure of tomato bushy stunt virus IV. The virus particle at 2.9-A  
resolution."  
RL J. MOL. BIOL. 171:61-93(1983).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RA HARRISON S.C., OLSON A.J., SCHUTT C.E., WINKLER F.K., BRICOGNE G.;  
RT "Tomato bushy stunt virus at 2.9-A resolution."  
RL NATURE 276:368-373(1978).  
CC -!- SUBUNIT: THE VIRUS COAT IS COMPOSED OF 180 COPIES OF THE COAT  
CC PROTEIN.  
CC -!- SIMILARITY: COAT PROTEINS FROM CARMOVIRUSES (CARMV, MNSV, TCV),  
CC DIANTHOVIRUSES (RCNMV), SOBEMOVIRUSES (NCMV, SBMV), TOMBUSVIRUSES  
CC (AMCV, CNV, CYRSV, AND TBSV), AND TNV ARE RELATED IN THEIR S  
CC REGION.  
DR PIR: S07259; S07259.  
DR PDB: 2TBV; 25-APR-86.  
DR PROSITE: PS00555; ICOSAH\_VIR\_COAT\_S; 1.  
DR PFAM: PF00729; Viral\_coat; 1.  
KW COAT PROTEIN; 3D-STRUCTURE.

FT DOMAIN 1 101 R DOMAIN, INTERACTION WITH RNA.  
FT DOMAIN 102 263 S DOMAIN, VIRION SHELL.  
FT DOMAIN 264 387 P DOMAIN, PROJECTING.  
FT STRAND 94 96  
FT TURN 100 101  
FT STRAND 105 107  
FT STRAND 109 116  
FT STRAND 121 123  
FT TURN 124 126  
FT TURN 139 140  
FT TURN 145 147  
FT TURN 148 150  
FT STRAND 154 155  
FT STRAND 157 165  
FT TURN 170 171  
FT STRAND 176 179  
FT TURN 184 185  
FT HELIX 192 195  
FT TURN 196 197  
FT TURN 199 200  
FT STRAND 202 204  
FT STRAND 210 213  
FT STRAND 221 222  
FT TURN 231 234  
FT STRAND 240 244  
FT STRAND 251 262  
FT STRAND 278 278  
FT STRAND 293 294  
FT STRAND 297 298  
FT TURN 299 300  
FT STRAND 301 302  
FT STRAND 304 306  
FT STRAND 310 319  
FT STRAND 325 330  
FT STRAND 333 340  
FT STRAND 347 355  
FT TURN 358 359  
FT STRAND 360 365  
FT STRAND 371 376  
SQ SEQUENCE 387 AA: 40540 MW: 98A2DBDE CRC32;  
Query Match 19.4%; Score 6; DB 1; Length 387;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 246 GAGADA 251  
QY 3 GAGADA 8  
RESULT 12  
ID AMY\_BACAM STANDARD; PRT: 514 AA.  
AC P00692;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN  
DE GLUCANOHYDROLASE).  
OS BACILLUS AMYLOLIQUEFACIENS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IH;  
RX MEDLINE: 83108080.  
RA TAKKINEN K., PETTERSSON R.F., KALKKINEN N., PALVA I., SODERLUND H.,  
RA KAARIAINEN L.;  
RT "Amino acid sequence of alpha-amylase from Bacillus amyloliquefaciens  
deduced from the nucleotide sequence of the cloned gene."  
RL J. BIOL. CHEM. 258:1007-1013(1983).  
RN [2]  
RP SEQUENCE OF 32-222.  
RX MEDLINE: 80241725.

RA CHUNG H.S., FRIEDBERG F.;  
RT "Sequence of the N-terminal half of *Bacillus amyloliquefaciens* alpha-  
RL amylase";  
RN BIOCHEM. J. 185:387-395(1980).  
RX [3]  
RY SEQUENCE OF 1-96 FROM N.A.  
RZ MEDLINE; 82051296.  
RA PALVA I., PETTERSSON R.F., KALKKINEN N., LEHTOVAARA P., SARVAS M.,  
RA SODERLUND H., TAKINEN K., KAARIJAINEN L.;  
RT "Nucleotide sequence of the promoter and NH2-terminal signal peptide  
RT region of the alpha-amylase gene from *Bacillus amyloliquefaciens*";  
RL GENE 15:43-51(1981).  
RN [4]  
RP SEQUENCE OF 1-39 FROM N.A.  
RX MEDLINE; 88137952.  
RA RUOHONEN L., HACKMAN P., LEHTOVAARA P., KNOWLES J.K.C., KARAENEN S.;  
RT "Efficient secretion of *Bacillus amyloliquefaciens* alpha-amylase by  
RT its own signal peptide from *Saccharomyces cerevisiae* host cells";  
RL GENE 59:161-170(1987).  
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: J01542; G142429; -  
DR EMBL: V00092; G39298; -  
DR EMBL: A20154; G580683; -  
DR EMBL: M18424; G142431; -  
DR PIR; A00843; ALBSN.  
DR FPM; PF00128; alpha-amylase; 1.  
DR HSPM; P06278; 1BPL.  
KW HYDROLASE; GLYCOSIDASE; CARBOHYDRATE METABOLISM; SIGNAL.  
FT SIGNAL 1 31  
FT CHAIN 32 514  
FT ACT\_SITE 262 514 ALPHA-AMYLASE.  
FT ACT\_SITE 262 262 BY SIMILARITY.  
FT ACT\_SITE 266 266 BY SIMILARITY.  
FT ACT\_SITE 359 359 BY SIMILARITY.  
FT CONFLICT 54 54 L -> I (IN REF. 2).  
FT CONFLICT 64 64 I -> L (IN REF. 2).  
FT CONFLICT 79 79 S -> D (IN REF. 2).  
FT CONFLICT 84 84 G -> S (IN REF. 2).  
SQ SEQUENCE 514 AA; 58403 MW; A37712F3 CRC32;  
  
Query Match 19.48; Score 6; DB 1; Length 514;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 136 AGADAT 141  
QY 4 AGADAT 9  
|||||  
  
RESULT 13  
ID SECD\_MYCLE STANDARD; PRT; 701 AA.  
AC P38387;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE PROTEIN-EXPORT MEMBRANE PROTEIN SECD.  
GN SECD OR B1177 CL164.  
OS MYCOBACTERIUM LEPRAE.  
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIACEAE;  
OC ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIUM.  
RN [1]

RP SEQUENCE FROM N.A.  
RA SMITH D.R., ROBISON K.;  
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: INVOLVED IN PROTEIN EXPORT (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: IN THE C-TERMINAL, WITH SECF.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U00011; G466811; -  
DR PROTEIN TRANSPORT; TRANSLOCATION; TRANSMEMBRANE; MEMBRANE.  
KW TRANSMEM 225 245 POTENTIAL.  
FT TRANSMEM 296 316 POTENTIAL.  
FT TRANSMEM 321 341 POTENTIAL.  
FT TRANSMEM 352 372 POTENTIAL.  
FT TRANSMEM 400 420 POTENTIAL.  
FT TRANSMEM 425 445 POTENTIAL.  
SQ SEQUENCE 701 AA; 75281 MW; 015A57AC CRC32;  
  
Query Match 19.48; Score 6; DB 1; Length 701;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 512 DNGAGA 517  
QY 1 DNGAGA 6  
|||||  
  
RESULT 14  
ID FHUA\_ECOLI STANDARD; PRT; 747 AA.  
AC P06971; P75665; P71280;  
DT 01-APR-1988 (REL. 07, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE FERRICHRONE-IRON RECEPTOR PRECURSOR (FERRIC HYDROXAMATE UPTAKE).  
GN FHUA OR TONA.  
OS ESCHERICHIA COLI.  
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
OC ESCHERICHIA.  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-47.  
RC STRAIN-K12;  
RX MEDLINE; 86085668.  
RA COULTON J.W., MASON P., CAMERON D.R., CARMEL G., JEAN R., RODE H.N.;  
RT "Protein fusions of beta-galactosidase to the ferrichrome-iron  
RT receptor of *Escherichia coli* K-12";  
RL J. BACTERIOL. 165:181-192(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / W3110;  
RX MEDLINE; 94261430.  
RA FUJITA N., MORI H., YURA T., ISHIHAMA A.;  
RT "Systematic sequencing of the *Escherichia coli* genome: analysis of  
RT the 2.4-4.1 min (110,917-193,643 bp) region";  
RL NUCLEIC ACIDS RES. 22:1637-1639(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE; 97426617.  
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,  
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,  
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,  
RA MAU B., SHAO Y.;  
RT "The complete genome sequence of *Escherichia coli* K-12";  
RL SCIENCE 277:1453-1474(1997).  
RN [4]  
RP SEQUENCE OF 482-647 FROM N.A.

RA SCHRAMM S., DUNCAN M., ALLEN E., ARAUJO R., APARICIO A., CHUNG E.,  
RA DAVIS K., FEDERSPIEL N., HYMAN R., KALMAN S., KOMP C., KURDI O.,  
RA LASHKARI D., LEW H., LIN D., NAMATH A., OEFNER P., ROBERTS D.,  
RA DAVIS R.W.,  
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [5]  
RP SEQUENCE OF 723-747 FROM N.A.  
RX MEDLINE: 88038363.  
RA BURKHARDT R., BRAUN V.,  
RT "Nucleotide sequence of the fhuc and fhud genes involved in iron  
RT (iii) hydroxamate transport: domains in FhuC homologous to  
RT ATP-binding proteins.";  
RL MOL. GEN. GENET. 209:49-55(1987).  
RN [6]  
RP REVIEW.  
RX MEDLINE: 94265918.  
RA BRAUN V., KILLMAN H., BENZ R.,  
RT "Energy-coupled transport through the outer membrane of Escherichia  
RT coli small deletions in the gating loop convert the FhuA transport  
RT protein into a diffusion channel.";  
RL FEBS LETT. 346:59-64(1994).  
RN [7]  
RP ION CHANNEL.  
RX MEDLINE: 96203106.  
RA BONHIVERS M., GHAZI A., BOULANGER P., LEPELLIER L.,  
RT "FhuA, a transporter of the Escherichia coli outer membrane, is  
RT converted into a channel upon binding of bacteriophage T5.";  
RL EMBO J. 15:1850-1856(1996).  
CC -!- FUNCTION: THIS RECEPTOR BINDS THE FERRIC-IRON LIGAND. IT  
CC INTERACTS WITH THE TONB PROTEIN, WHICH IS RESPONSIBLE FOR ENERGY  
CC COUPLING OF THE FERRIC-IRON-PROMOTED IRON TRANSPORT SYSTEM. ACTS  
CC AS A RECEPTOR FOR BACTERIOPHAGE T5 AS WELL AS T1, PH180 AND  
CC COLICIN M. BINDING OF T5 TRIGGERS THE OPENING OF A HIGH  
CC CONDUCTANCE ION CHANNEL.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. OUTER MEMBRANE.  
CC -!- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M12486; G2226438; -.  
DR EMBL: D26562; G473809; -.  
DR EMBL: AE000124; G1786344; -.  
DR EMBL: U70214; G1552728; -.  
DR EMBL: X05810; G41439; -.  
DR PIR: A25196; QRECFE.  
DR PIR: S06358; S06358.  
DR PIR: S45219; S45219.  
DR ECOGENE: EG10302; FHUA.  
DR PROSITE: PS00430; TONB\_DEPENDENT\_REC\_1; 1.  
DR PROSITE: PS01156; TONB\_DEPENDENT\_REC\_2; 1.  
DR PFAM: PF00593; TonB\_boxC; 1.  
KW OUTER MEMBRANE; IRON TRANSPORT; TRANSPORT; TONB BOX; SIGNAL;  
KW PHAGE RECOGNITION; RECEPTOR; TRANSMEMBRANE.  
FT SIGNAL 1 33  
FT CHAIN 34 747 FERRIC-IRON RECEPTOR.  
FT SIMILAR 40 47 TONB BOX.  
FT SIMILAR 730 747 TONB C-TERMINAL BOX.  
FT CONFLICT 609 610 AA -> RP (IN REF. 2).  
FT CONFLICT 737 737 R -> P (IN REF. 5).  
SQ SEQUENCE 747 AA; 82182 MW; 5F4EF63B CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 747;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 219 DDGVYS 224  
|||||

QY 11 DDGVYS 16  
  
RESULT 15  
ID FLO8\_YEAST STANDARD; PRY: 799 AA.  
AC P40068; P40067; Q05751;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE TRANSCRIPTIONAL ACTIVATOR FLO8 (PDH5 PROTEIN).  
GN FLO8 OR PDH5 OR YER109C/YER108C.  
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;  
OC SACHAROMYCETACEAE; SACHAROMYCES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SIGMA 1278B;  
RA LIU H., STYLES C.A., FINK G.;  
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE OF 71-799 FROM N.A.  
RC STRAIN-ATCC 60715;  
RX MEDLINE: 96335146.  
RA KOBAYASHI O., SUDA H., OHTANI T., SONE H.;  
RT "Molecular cloning and analysis of the dominant flocculation gene  
RT FLO8 from Saccharomyces cerevisiae.";  
RL MOL. GEN. GENET. 251:707-715(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C / AB972;  
RA DIETRICH F.S., MULLIGAN J.T., HENNESSEY K.M., ALLEN E., ARAUJO R.,  
RA AVILES E., BERNO A., BRENNAN T., CARPENTER J., CHEN E., CHERRY J.M.,  
RA CHUNG E., DUNCAN M., GUZMAN E., HARTZELL G., HUNICKE-SMITH S.,  
RA HYMAN R., KAYSER A., KOMP C., LASHKARI D., LEW H., LIN D.,  
RA MOSEDALE D., NAKAHARA K., NAMATH A., NORGREN R., OEFNER P., OH C.,  
RA PETEL F.X., ROBERTS D., SEHL P., SCHRAMM S., SHOGEN T., SMITH V.,  
RA TAYLOR P., WEI Y., YELTON M., BOTSTEIN D., DAVIS R.W.;  
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: REQUIRED FOR DIPLOID FILAMENTOUS GROWTH. HAPLOID  
CC INVASIVE GROWTH AND FLOCCULATION. PUTATIVE TRANSCRIPTIONAL  
CC ACTIVATOR OF FLO1.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: U51431; G1256780; -.  
DR EMBL: D83713; G1213591; -.  
DR EMBL: U18916; G603348; ALT\_FRAME.  
DR EMBL: U18916; G603347; ALT\_FRAME.  
DR SGD: L0003202; FLO8.  
KW TRANSCRIPTION REGULATION; ACTIVATOR; NUCLEAR PROTEIN.  
FT DOMAIN 41 55 POLY-GLN.  
FT CONFLICT 112 112 V -> I (IN REF. 2).  
FT CONFLICT 115 115 P -> S (IN REF. 2).  
FT CONFLICT 383 383 C -> G (IN REF. 2).  
FT CONFLICT 441 441 A -> T (IN REF. 2).  
FT CONFLICT 447 447 A -> V (IN REF. 2).  
FT CONFLICT 598 598 R -> P (IN REF. 2).  
SQ SEQUENCE 799 AA; 86834 MW; C62DF080 CRC32;  
  
Query Match 19.4%; Score 6; DB 1; Length 799;  
Best Local Similarity 100.0%; Pred. No. 4.59e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 505 ADATKD 510

Qy       |||||  
         6 ADATKD 11

Search completed: Sat Aug 28 15:01:29 1999  
Job time : 42 secs.

This Page Blank (uspto)

\*\*\*\*\*  
W A S E R E H  
\*\*\*\*\* (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:01:46 1999; MasPar time 7.66 Seconds  
Tabular output not generated. 220.813 Million cell updates/sec

Title: >US-09-049-696-47  
Description: (1-31) from US09049696.pep  
Perfect Score: 31  
Sequence: 1 DNGAGADATKDCGVSYRYFTTYDNGRYSVK 31

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: sptrembl9

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organella  
9:sp-phase 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 2.681; Variance 0.433; scale 6.188

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	10	32.3	913 11	088826	G0B-5 PROTEIN.	4.08e-07
2	7	22.6	326 2	050269	HYPOTHETICAL 35.8 KD P	2.09e-01
3	7	22.6	327 2	025525	GMP REDUCTASE (GUAC).	2.09e-01
4	7	22.6	483 8	036097	CYTCHROME OXIDASE SUB	2.09e-01
5	7	22.6	901 11	088860	CHLORIDE CHANNEL CACC.	2.09e-01
6	7	22.6	4910 3	Q12019	SIMILARITY NEAR N-TERM	2.09e-01
7	6	19.4	22 10	Q39606	NIT-1 GENE ENCODING NI	1.00e+01
8	6	19.4	97 2	Q46404	CORRESPONDS TO A 97 AM	1.00e+01
9	6	19.4	104 11	Q63144	OUTER MEMBRANE CYTOCHR	1.00e+01
10	6	19.4	158 5	Q02430	CYTOSOLIC CU-ZN SUPERO	1.00e+01
11	6	19.4	166 3	Q08004	CHROMOSOME XII READING	1.00e+01
12	6	19.4	171 8	047877	NADH DEHYDROGENASE SUB	1.00e+01
13	6	19.4	183 2	Q49286	MGPA (FRAGMENT).	1.00e+01
14	6	19.4	186 2	Q86545	HYPOTHETICAL 19.9 KD P	1.00e+01
15	6	19.4	197 10	Q65731	40S RIBOSOMAL PROTEIN	1.00e+01
16	6	19.4	212 2	O50342	TYPE B PILLIN GENE (STR	1.00e+01
17	6	19.4	272 2	O06129	TRPC.	1.00e+01
18	6	19.4	278 2	O86901	HYDROLASE, DXNB PROTEI	1.00e+01
19	6	19.4	309 2	O68530	ENDONUCLEASE S1 HOMOLO	1.00e+01
20	6	19.4	318 10	Q05957	PHOSPHONATE BIOSYNTHES	1.00e+01

21	6	19.4	343 2	Q55989	HYPOTHETICAL 38.4 KD P	1.00e+01
22	6	19.4	350 10	O81850	PUTATIVE FE(II) TRANSP	1.00e+01
23	6	19.4	357 10	Q24927	NARINGENIN 3-DIOXYGENA	1.00e+01
24	6	19.4	366 5	Q24830	CNPGO (FRAGMENT).	1.00e+01
25	6	19.4	375 5	O18678	G PROTEIN ALPHA SUBUNI	1.00e+01
26	6	19.4	385 5	O44744	R06A10.2 PROTEIN.	1.00e+01
27	6	19.4	386 13	O73860	OVALBUMIN.	1.00e+01
28	6	19.4	397 2	O85687	HOOK ASSOCIATED PROTEI	1.00e+01
29	6	19.4	400 2	O31200	IMIDAZOLONE PROPIONATE	1.00e+01
30	6	19.4	401 5	O76429	TRANSFERRIN-BINDING PR	1.00e+01
31	6	19.4	402 2	O69865	PUTATIVE LYSINE/ORNITH	1.00e+01
32	6	19.4	414 14	P89954	CORE ENV AND PART OF	1.00e+01
33	6	19.4	471 2	O69486	PROTEIN EXPORT MEMBRAN	1.00e+01
34	6	19.4	484 5	Q19053	E04D5.3 PROTEIN.	1.00e+01
35	6	19.4	506 2	Q53834	PHASE-2 FLAGELLIN STRU	1.00e+01
36	6	19.4	534 2	Q54707	IMMUNOGENIC SECRETED P	1.00e+01
37	6	19.4	536 5	O15782	CHAPERONIN 60.	1.00e+01
38	6	19.4	548 2	O05404	HYPOTHETICAL 55.1 KD P	1.00e+01
39	6	19.4	566 5	Q21740	R05D11.8 PROTEIN.	1.00e+01
40	6	19.4	569 1	O27599	TUNGSTEN FORMYL METHANO	1.00e+01
41	6	19.4	636 2	O88039	PUTATIVE ABC TRANSPORT	1.00e+01
42	6	19.4	636 2	Q53818	POTENTIAL ATP-BINDING	1.00e+01
43	6	19.4	655 5	O24937	TRANSHYDROGENASE (FRAG	1.00e+01
44	6	19.4	732 2	O86424	FERRICHRONE IRON RECEP	1.00e+01
45	6	19.4	794 6	O18742	LU-ECAM-1.	1.00e+01
46	6	19.4	820 6	O18743	LU-ECAM-1.	1.00e+01
47	6	19.4	836 5	O62201	F32B4.8 PROTEIN.	1.00e+01
48	6	19.4	905 6	O18741	LU-ECAM-1.	1.00e+01
49	6	19.4	982 2	O54142	POLYKETIDE SYNTHASE RE	1.00e+01
50	6	19.4	1023 5	O07600	NAD(P) TRANSHYDROGENAS	1.00e+01
51	6	19.4	1099 2	O54377	EXONUCLEASE REXB.	1.00e+01
52	6	19.4	1489 2	O53559	PGRS-FAMILY PROTEIN.	1.00e+01
53	6	19.4	1643 10	O81486	MYOSIN HEAVY CHAIN.	1.00e+01
54	6	19.4	2533 5	P90589	ALPHA-51D IMMOBILIZATI	1.00e+01
55	6	19.4	2533 5	O27183	ALPHA-51D-IMMOBILIZATI	1.00e+01
56	6	19.4	2543 5	P90649	156D SUFACE ANTIGEN.	1.00e+01
57	6	19.4	3015 14	O92532	POLIPROTEIN.	1.00e+01
58	6	19.4	15281 3	Q09164	CYCLOSOPORIN SYNTHETASE	1.00e+01
59	5	16.1	26 14	O82944	VP-1 (FRAGMENT).	3.25e+02
60	5	16.1	26 14	O82943	VP-1 (FRAGMENT).	3.25e+02
61	5	16.1	65 10	O43667	TRICHOSANTHES TRYPSIN	3.25e+02
62	5	16.1	66 14	Q68561	ENVELOPE GLYCOPROTEIN	3.25e+02
63	5	16.1	67 14	O84487	GENOME, PARTIAL SEQUEN	3.25e+02
64	5	16.1	70 6	P79261	GLUTAMIC ACID DECARBOX	3.25e+02
65	5	16.1	72 3	Q00411	CHITIN SYNTHASE (FRAGM	3.25e+02
66	5	16.1	79 2	O07901	MINOR OUTER MEMBRANE P	3.25e+02
67	5	16.1	81 1	O28454	HYPOTHETICAL 9.5 KD PR	3.25e+02
68	5	16.1	89 7	O62865	MAJOR HISTOCOMPATIBILI	3.25e+02
69	5	16.1	89 7	O62866	MAJOR HISTOCOMPATIBILI	3.25e+02
70	5	16.1	89 7	O62866	MAJOR HISTOCOMPATIBILI	3.25e+02
71	5	16.1	89 2	O54175	PUTATIVE ACYL-CARRIER	3.25e+02
72	5	16.1	89 2	O54219	ACYL CARRIER PROTEIN.	3.25e+02
73	5	16.1	92 2	O06410	HYPOTHETICAL 9.7 KD PR	3.25e+02
74	5	16.1	93 3	Q12497	CHROMOSOME-IV READING	3.25e+02
75	5	16.1	95 2	Q14914	DEHYDROGENASE (FRAGMEN	3.25e+02
76	5	16.1	97 14	O88972	RNA-DEPENDENT RNA POLY	3.25e+02
77	5	16.1	98 2	O84927	CSOSIA.	3.25e+02
78	5	16.1	98 2	O84926	CSOSIC.	3.25e+02
79	5	16.1	99 2	O84922	CARBOXYISOME SHELL PEPT	3.25e+02
80	5	16.1	101 2	O84922	CARBOXYISOME SHELL PEPT	3.25e+02
81	5	16.1	102 5	O62385	T26H5.4 PROTEIN.	3.25e+02
82	5	16.1	102 14	O88976	RNA-DEPENDENT RNA POLY	3.25e+02
83	5	16.1	102 14	O88975	RNA-DEPENDENT RNA POLY	3.25e+02
84	5	16.1	105 14	Q98584	PBCV-1 UNIDENTIFIED OR	3.25e+02
85	5	16.1	108 2	O84928	CSOSIB.	3.25e+02
86	5	16.1	109 2	O51031	VARIANT P9 PILE (FRAGM	3.25e+02
87	5	16.1	111 5	O46153	CATHEPSIN L2 (FRAGMENT	3.25e+02
88	5	16.1	113 4	O60739	GC20 PROTEIN.	3.25e+02
89	5	16.1	118 2	O54189	PUATIVE CORALT TRANSPO	3.25e+02
90	5	16.1	118 5	O27454	MICROFILARIAL CHITININ	3.25e+02
91	5	16.1	119 6	O77517	BETA-2-MICROGLOBULIN P	3.25e+02
92	5	16.1	119 6	O77516	BETA-2-MICROGLOBULIN P	3.25e+02
93	5	16.1	122 2	O84923	CARBOXYISOME SHELL PEPT	3.25e+02

94	5	16.1	123	10	Q42253	LEA PROTEIN (FRAGMENT)	3.25e+02	167	183	14	041272	ORF 552.	3.25e+02
95	5	16.1	124	2	Q54238	ORF10.	3.25e+02	168	183	2	Q52885	ORF9.	3.25e+02
96	5	16.1	124	14	Q88974	RNA-DEPENDENT RNA POLY	3.25e+02	169	183	11	Q63740	PDGF PROTEIN (FRAGMENT	3.25e+02
97	5	16.1	127	5	Q16413	C3G5.1 PROTEIN.	3.25e+02	170	184	5	Q16042	ANON1A4.	3.25e+02
98	5	16.1	127	5	Q16908	CYTCHROME P450 (FRAGM	3.25e+02	171	184	2	Q52091	OUTER SURFACE PROTEIN	3.25e+02
99	5	16.1	134	2	Q95485	ANKYRIN (FRAGMENT).	3.25e+02	172	187	2	Q07394	PUTATIVE IRON-SULFUR P	3.25e+02
100	5	16.1	136	2	Q87897	DSRF.	3.25e+02	173	188	5	Q25352	L2759.14.	3.25e+02
101	5	16.1	137	10	Q04572	T7N9.25.	3.25e+02	174	188	3	Q14284	HYPOTHETICAL 21.5 KD P	3.25e+02
102	5	16.1	138	14	Q88186	RNA-DEPENDENT RNA POLY	3.25e+02	175	190	10	Q40995	ALBUMIN 1.	3.25e+02
103	5	16.1	138	1	Q46209	ILES GENE (FRAGMENT).	3.25e+02	176	190	2	Q06882	PUTATIVE VLS RECOMBINA	3.25e+02
104	5	16.1	138	2	Q53318	ORF1 5'OF ATWA (FRAGME	3.25e+02	177	192	14	Q89660	(HHV-6).	3.25e+02
105	5	16.1	139	14	Q88185	RNA-DEPENDENT RNA POLY	3.25e+02	178	192	2	Q78287	GLUTAMINE SYNTHETASE (	3.25e+02
106	5	16.1	140	2	Q83445	V-TYPE ATPASE, SUBUNIT	3.25e+02	179	194	2	Q05304	HYPOTHETICAL 19.9 KD P	3.25e+02
107	5	16.1	143	2	Q33974	VACUOLATING TOXIN VACA	3.25e+02	180	194	10	Q81607	1-AMINOCYCLOPROPANE-1-	3.25e+02
108	5	16.1	145	2	Q33975	VACUOLATING TOXIN VACA	3.25e+02	181	194	2	Q94247	OUTER SURFACE PROTEIN	3.25e+02
109	5	16.1	145	14	Q88188	RNA-DEPENDENT RNA POLY	3.25e+02	182	195	2	Q54412	66 GLYT TRANSFER RNA-G	3.25e+02
110	5	16.1	146	2	Q06153	CYTCHROMICAL 15.3 KD P	3.25e+02	183	197	14	Q80837	COAT PROTEIN.	3.25e+02
111	5	16.1	147	2	Q89087	HYPOTHETICAL 15.3 KD P	3.25e+02	184	198	2	Q75474	ARGININE DEIMINASE.	3.25e+02
112	5	16.1	147	2	Q89087	PUTATIVE REGULATOR.	3.25e+02	185	198	2	Q75474	ARGININE DEIMINASE.	3.25e+02
113	5	16.1	149	14	Q09787	RNA DEPENDENT RNA POLY	3.25e+02	186	199	2	Q31701	MOLYBDOPTEIN-GUANINE	3.25e+02
114	5	16.1	149	14	Q09784	RNA DEPENDENT RNA POLY	3.25e+02	187	199	2	Q95645	MOB GENE.	3.25e+02
115	5	16.1	149	14	Q09784	RNA DEPENDENT RNA POLY	3.25e+02	188	201	10	Q43325	KUNITZ PROTEASE INHIBI	3.25e+02
116	5	16.1	150	14	Q09736	RNA DEPENDENT RNA POLY	3.25e+02	189	201	8	Q96123	SUPEROXIDE DISMUTASE (	3.25e+02
117	5	16.1	150	14	Q09736	RNA DEPENDENT RNA POLY	3.25e+02	190	201	8	Q96123	SUPEROXIDE DISMUTASE (	3.25e+02
118	5	16.1	150	14	Q09789	RNA DEPENDENT RNA POLY	3.25e+02	191	202	2	Q54157	HYPOTHETICAL 22.2 KD P	3.25e+02
119	5	16.1	150	14	Q09786	RNA DEPENDENT RNA POLY	3.25e+02	192	202	2	Q50626	OUTER SURFACE PROTEIN	3.25e+02
120	5	16.1	150	14	Q09785	RNA DEPENDENT RNA POLY	3.25e+02	193	203	2	Q50619	OUTER SURFACE PROTEIN	3.25e+02
121	5	16.1	150	14	Q87622	ORF 1742.	3.25e+02	194	203	2	Q50624	OUTER SURFACE PROTEIN	3.25e+02
122	5	16.1	150	14	Q09776	RNA DEPENDENT RNA POLY	3.25e+02	195	205	3	Q12025	CHROMOSOME IV READING	3.25e+02
123	5	16.1	150	14	Q09788	RNA DEPENDENT RNA POLY	3.25e+02	196	209	2	Q07056	VMP-LIKE SEQUENCE PROT	3.25e+02
124	5	16.1	151	2	Q30732	CCOH.	3.25e+02	197	209	1	Q07117	HYPOTHETICAL 23.5 KD P	3.25e+02
125	5	16.1	156	2	Q04959	GLUTAMINE SYNTHETASE (	3.25e+02	198	210	2	Q07054	VMP-LIKE SEQUENCE PROT	3.25e+02
126	5	16.1	156	2	Q07486	HYPOTHETICAL 16.5 KD P	3.25e+02	199	211	2	Q07052	VMP-LIKE SEQUENCE PROT	3.25e+02
127	5	16.1	156	5	Q77302	40S RIBOSOMAL PROTEIN	3.25e+02	200	211	2	Q68351	VMP-LIKE SEQUENCE PROT	3.25e+02
128	5	16.1	158	10	Q43759	GLUTAMINE SYNTHETASE (	3.25e+02	201	212	2	Q68342	VMP-LIKE SEQUENCE PROT	3.25e+02
129	5	16.1	159	10	Q04960	GLUTAMINE SYNTHETASE (	3.25e+02	202	212	2	Q54626	DOT1.	3.25e+02
130	5	16.1	160	14	Q85300	HOMOLOGUE OF RETROVIRA	3.25e+02	203	212	2	Q68351	VMP-LIKE SEQUENCE PROT	3.25e+02
131	5	16.1	161	5	Q21525	M05D6.6 PROTEIN.	3.25e+02	204	212	2	Q07052	VMP-LIKE SEQUENCE PROT	3.25e+02
132	5	16.1	163	14	Q84686	RNA-DEPENDENT RNA POLY	3.25e+02	205	213	2	Q07052	VMP-LIKE SEQUENCE PROT	3.25e+02
133	5	16.1	163	11	Q35941	RING FINGER PROTEIN 4	3.25e+02	206	213	2	Q68354	VMP-LIKE SEQUENCE PROT	3.25e+02
134	5	16.1	164	10	Q82516	SOLUBLE ACID INVERTASE	3.25e+02	207	214	10	Q40116	RIBONUCLEASE (RNASE LC	3.25e+02
135	5	16.1	165	2	Q44025	HOXW PROTEIN.	3.25e+02	208	217	14	Q92405	ACMPV ORF38	3.25e+02
136	5	16.1	165	2	Q72308	HOXW.	3.25e+02	209	217	1	Q30641	MONOMETHYLAMINE CORBIN	3.25e+02
137	5	16.1	166	2	Q06370	HYPOTHETICAL 18.7 KD P	3.25e+02	210	219	10	Q41817	HEAT SHOCK PROTEIN 70	3.25e+02
138	5	16.1	167	10	Q04957	GLUTAMINE SYNTHETASE (	3.25e+02	211	219	5	Q20283	COSMID F41F3.	3.25e+02
139	5	16.1	170	2	Q87500	ORF19.	3.25e+02	212	219	1	Q52017	ORF H1354.	3.25e+02
140	5	16.1	171	10	Q40996	ALBUMIN 2 (FRAGMENT).	3.25e+02	213	219	1	Q28237	URIDYLATE KINASE (PYRH	3.25e+02
141	5	16.1	173	14	Q88174	POLYMERASE (FRAGMENT).	3.25e+02	214	220	2	Q54372	(78-11).	3.25e+02
142	5	16.1	173	14	Q88965	POLYMERASE (FRAGMENT).	3.25e+02	215	221	10	Q96477	LRR PROTEIN.	3.25e+02
143	5	16.1	173	14	Q87003	POLYMERASE (FRAGMENT).	3.25e+02	216	222	2	Q74298	REGULATORY COMPONENTS	3.25e+02
144	5	16.1	173	14	Q12372	RNA POLYMERASE (FRAGME	3.25e+02	217	222	2	Q56620	GALACTOSYL TRANSFERASE	3.25e+02
145	5	16.1	173	14	Q88973	RNA-DEPENDENT RNA POLY	3.25e+02	218	223	11	Q63211	GLUTAMATE DECARBOXYLAS	3.25e+02
146	5	16.1	173	14	Q87018	RNA-DEPENDENT RNA POLY	3.25e+02	219	225	2	Q53149	LIPOTEIN PRECURSOR	3.25e+02
147	5	16.1	173	10	Q42874	GLUTAMINE SYNTHETASE (	3.25e+02	220	228	2	Q59400	COLI FUMC FUMARASE (EC	3.25e+02
148	5	16.1	173	14	Q88183	RNA-DEPENDENT RNA POLY	3.25e+02	221	228	1	Q26232	OROTIDINE 5' MONOPHOSP	3.25e+02
149	5	16.1	173	14	Q87007	POLYMERASE (FRAGMENT).	3.25e+02	222	229	5	Q21069	KOIA11.2 PROTEIN.	3.25e+02
150	5	16.1	173	14	Q87005	POLYMERASE (FRAGMENT).	3.25e+02	223	230	14	Q90239	CAPSID PROTEIN.	3.25e+02
151	5	16.1	173	14	Q65398	RNA-DEPENDENT RNA POLY	3.25e+02	224	233	14	Q91868	ORF2.	3.25e+02
152	5	16.1	173	14	Q88176	POLYMERASE (FRAGMENT).	3.25e+02	225	233	14	Q91865	ORF2.	3.25e+02
153	5	16.1	174	2	Q05997	PUTATIVE FIBRIL ASSO	3.25e+02	226	233	14	Q56129	ORF2.	3.25e+02
154	5	16.1	174	2	Q95086	HYPOTHETICAL 17.8 KD P	3.25e+02	227	236	1	Q29732	STRAIN PMWS PCV, COMPL	3.25e+02
155	5	16.1	174	2	Q68798	HYPOTHETICAL 18.1 KD P	3.25e+02	228	236	13	Q42570	CONSERVED HYPOTHETICAL	3.25e+02
156	5	16.1	176	14	Q98179	MC008L.	3.25e+02	229	236	14	Q66680	HOMEBOX PROTEIN HOXC-	3.25e+02
157	5	16.1	178	5	Q44809	FL4D2.2 PROTEIN.	3.25e+02	230	241	1	Q58462	VIORIN PROTEIN US10.	3.25e+02
158	5	16.1	179	2	Q44981	OUTER SURFACE PROTEIN	3.25e+02	231	241	1	Q96737	241AA LONG HYPOTHETICA	3.25e+02
159	5	16.1	179	2	Q44981	K08C7.4 PROTEIN.	3.25e+02	232	242	14	Q96737	PVIII.	3.25e+02
160	5	16.1	180	7	Q79562	HUMAN LEUKOCYTE ANTIGE	3.25e+02	233	243	5	Q22341	SIMILAR TO C. ELEGANS	3.25e+02
161	5	16.1	180	5	Q91127	CODED FOR BY C. ELEGAN	3.25e+02	234	243	5	Q17587	COLH6.8 PROTEIN.	3.25e+02
162	5	16.1	181	7	Q62892	MHC CLASS I ANTIGEN (F	3.25e+02	235	244	2	Q03074	TRANSMEMBRANE PROTEIN	3.25e+02
163	5	16.1	181	7	Q19521	MHC CLASS I ANTIGEN HL	3.25e+02	236	244	14	Q96729	PVIII, FIBER, ORF-8, O	3.25e+02
164	5	16.1	181	14	Q88187	RNA-DEPENDENT RNA POLY	3.25e+02	237	244	10	Q80327	THAUMATIN-LIKE PROTEIN	3.25e+02
165	5	16.1	181	7	Q62924	MHC CLASS I ANTIGEN (F	3.25e+02	238	245	10	Q22816	POLYGLACTURONASE ISOL	3.25e+02
166	5	16.1	182	1	Q74068	TATA BOX BINDING PROTE	3.25e+02	239	245	10	Q23478	SIMILARITY TO NARINGEN	3.25e+02
											082546	THAUMATIN-LIKE PROTEIN	3.25e+02



240	5	16.1	5	246	2	006711	3	25e+02	313	5	16.1	327	2	052907	SIALIC ACID SYNTHASE (	3	25e+02
241	5	16.1	5	247	10	022902	3	25e+02	314	5	16.1	330	10	004482	ESTS GB145673..	3	25e+02
242	5	16.1	5	248	2	095555	3	25e+02	315	5	16.1	330	2	054308	HELIIX-TURN-HELIX DNA B	3	25e+02
243	5	16.1	5	249	5	044115	3	25e+02	316	5	16.1	331	5	046032	CYSTEINE PROTEINASE.	3	25e+02
244	5	16.1	5	248	1	026652	3	25e+02	317	5	16.1	331	2	007608	HYPOTHETICAL 38.0 KD P	3	25e+02
245	5	16.1	5	249	1	028028	3	25e+02	318	5	16.1	331	2	048592	LLAI.1	3	25e+02
246	5	16.1	5	249	2	051641	3	25e+02	319	5	16.1	331	5	046031	CYSTEINE PROTEINASE.	3	25e+02
247	5	16.1	5	251	5	017435	3	25e+02	320	5	16.1	331	5	017151	C41H7.4 PROTEIN.	3	25e+02
248	5	16.1	5	251	2	066154	3	25e+02	321	5	16.1	332	3	002796	LPE17P.	3	25e+02
249	5	16.1	5	253	1	058647	3	25e+02	322	5	16.1	333	2	053785	O-SUCCINYL BENZOIC ACID	3	25e+02
250	5	16.1	5	254	3	001136	3	25e+02	323	5	16.1	333	2	073635	PERIPLASMIC PHOSPHATE	3	25e+02
251	5	16.1	5	255	2	068049	3	25e+02	324	5	16.1	333	10	080408	PHARBITIS KNUTED-LIKE	3	25e+02
252	5	16.1	5	256	4	014908	3	25e+02	325	5	16.1	335	10	049290	PURATIVE CARBOXYPHOSH	3	25e+02
253	5	16.1	5	258	10	049194	3	25e+02	326	5	16.1	336	2	025910	QUINOLINATE SYNTHETASE	3	25e+02
254	5	16.1	5	259	14	092452	3	25e+02	327	5	16.1	337	2	067815	PHOSPHATE-BINDING PERI	3	25e+02
255	5	16.1	5	259	2	007198	3	25e+02	328	5	16.1	337	5	001592	SIMILAR TO CYSTEINE SY	3	25e+02
256	5	16.1	5	260	2	050504	3	25e+02	329	5	16.1	338	2	07061	CCAR AND BLP GENES AND	3	25e+02
257	5	16.1	5	260	2	032208	3	25e+02	330	5	16.1	338	5	046030	CYSTEINE PROTEINASE.	3	25e+02
258	5	16.1	5	261	3	000344	3	25e+02	331	5	16.1	338	2	085729	HYPOTHETICAL 36.6 KD P	3	25e+02
259	5	16.1	5	261	1	027772	3	25e+02	332	5	16.1	339	2	025067	ALIPHATIC AMIDASE (AIM	3	25e+02
260	5	16.1	5	262	2	005826	3	25e+02	333	5	16.1	339	2	032644	AMIE PROTEIN, ALIPHATI	3	25e+02
261	5	16.1	5	262	10	022973	3	25e+02	334	5	16.1	340	9	037851	CAPSID PORTAL PROTEIN.	3	25e+02
262	5	16.1	5	263	14	089773	3	25e+02	335	5	16.1	341	3	000057	PLANTA INDUCED RUST PR	3	25e+02
263	5	16.1	5	263	5	002570	3	25e+02	336	5	16.1	342	10	043065	GLUTAMINE SYNTHETASE (	3	25e+02
264	5	16.1	5	265	2	066687	3	25e+02	337	5	16.1	342	3	078690	PUTATIVE ATPASE (FRAGM	3	25e+02
265	5	16.1	5	266	5	024091	3	25e+02	338	5	16.1	343	5	090840	SIMILAR TO RAS-LIKE PR	3	25e+02
266	5	16.1	5	271	10	039960	3	25e+02	339	5	16.1	343	2	072467	ORF4.	3	25e+02
267	5	16.1	5	271	10	022740	3	25e+02	340	5	16.1	343	5	023320	2C443.1 PROTEIN.	3	25e+02
268	5	16.1	5	272	2	066527	3	25e+02	341	5	16.1	344	11	061893	HISTOCOMPATIBILITY 2,	3	25e+02
269	5	16.1	5	272	5	077216	3	25e+02	342	5	16.1	347	10	043242	PLASMA MEMBRANE H+ ATP	3	25e+02
270	5	16.1	5	273	2	071736	3	25e+02	343	5	16.1	347	5	021702	R04B5.5 PROTEIN.	3	25e+02
271	5	16.1	5	274	5	016133	3	25e+02	344	5	16.1	347	4	075227	R31449_1 (FRAGMENT).	3	25e+02
272	5	16.1	5	274	5	017086	3	25e+02	345	5	16.1	347	10	043241	PLASMA-MEMBRANE H+ ATP	3	25e+02
273	5	16.1	5	275	2	050711	3	25e+02	346	5	16.1	347	5	024502	SYNAPSE-ASSOCIATED PRO	3	25e+02
274	5	16.1	5	278	2	054227	3	25e+02	347	5	16.1	347	5	021703	R04B5.6 PROTEIN.	3	25e+02
275	5	16.1	5	279	2	054227	3	25e+02	348	5	16.1	348	3	074428	HYPOTHETICAL 37.9 KD P	3	25e+02
276	5	16.1	5	281	5	046137	3	25e+02	349	5	16.1	349	10	020305	GLUTAMINE SYNTHETASE (	3	25e+02
277	5	16.1	5	282	5	025395	3	25e+02	350	5	16.1	350	5	020305	COSMID F41H10.	3	25e+02
278	5	16.1	5	282	5	003432	3	25e+02	351	5	16.1	350	2	032290	XYAL PROTEIN.	3	25e+02
279	5	16.1	5	283	5	061141	3	25e+02	352	5	16.1	351	2	071747	HYPOTHETICAL 37.4 KD P	3	25e+02
280	5	16.1	5	283	1	058665	3	25e+02	353	5	16.1	351	3	078825	FISSION YEAST (FRAGMEN	3	25e+02
281	5	16.1	5	283	1	026892	3	25e+02	354	5	16.1	351	3	078825	SYNAPSE ASSOCIATED PRO	3	25e+02
282	5	16.1	5	285	2	053410	3	25e+02	355	5	16.1	352	10	022504	CYTOSOLIC GLUTAMINE SY	3	25e+02
283	5	16.1	5	286	2	084538	3	25e+02	356	5	16.1	352	10	024153	S-ADENOSYL-METHIONINE-	3	25e+02
284	5	16.1	5	287	2	086719	3	25e+02	357	5	16.1	353	5	091233	COSMID F08D12.	3	25e+02
285	5	16.1	5	287	10	023403	3	25e+02	358	5	16.1	353	2	097033	VARIABLE MAJOR PROTEIN	3	25e+02
286	5	16.1	5	288	14	084565	3	25e+02	359	5	16.1	354	14	082949	VP1 CAPSID PROTEIN.	3	25e+02
287	5	16.1	5	290	5	019574	3	25e+02	360	5	16.1	354	14	040909	VARIANT GS/B, COMPLETE	3	25e+02
288	5	16.1	5	293	2	083872	3	25e+02	361	5	16.1	354	14	058666	MAJOR CAPSID PROTEIN.	3	25e+02
289	5	16.1	5	294	4	092493	3	25e+02	362	5	16.1	354	10	024333	GLUTAMINE SYNTHETASE (	3	25e+02
290	5	16.1	5	298	10	040839	3	25e+02	363	5	16.1	354	5	016003	COLLAGEN-LIKE PROTEIN	3	25e+02
291	5	16.1	5	299	2	049149	3	25e+02	364	5	16.1	354	14	089547	VP1 (CAPSID PROTEIN).	3	25e+02
292	5	16.1	5	299	4	075636	3	25e+02	365	5	16.1	355	10	043760	GLUTAMINE SYNTHETASE (	3	25e+02
293	5	16.1	5	305	2	068121	3	25e+02	366	5	16.1	356	14	092786	VP1 MAJOR CAPSID PROTE	3	25e+02
294	5	16.1	5	309	2	052714	3	25e+02	367	5	16.1	356	2	071744	HYPOTHETICAL 37.4 KD P	3	25e+02
295	5	16.1	5	310	2	045414	3	25e+02	368	5	16.1	356	10	042623	GLUTAMINE SYNTHETASE (	3	25e+02
296	5	16.1	5	310	2	056167	3	25e+02	369	5	16.1	356	10	004880	GLUTAMINE SYNTHETASE (	3	25e+02
297	5	16.1	5	312	3	014047	3	25e+02	370	5	16.1	356	10	004852	GLUTAMINE SYNTHETASE (	3	25e+02
298	5	16.1	5	312	3	018097	3	25e+02	371	5	16.1	356	10	082560	GLUTAMINE SYNTHETASE (	3	25e+02
299	5	16.1	5	313	1	028287	3	25e+02	372	5	16.1	356	10	042899	GLUTAMINE SYNTHETASE (	3	25e+02
300	5	16.1	5	314	10	065030	3	25e+02	373	5	16.1	356	10	022331	GLUTAMINE SYNTHETASE (	3	25e+02
301	5	16.1	5	314	2	056167	3	25e+02	374	5	16.1	356	10	004998	GLUTAMINE SYNTHETASE (	3	25e+02
302	5	16.1	5	317	10	023085	3	25e+02	375	5	16.1	356	10	042950	GLUTAMINE SYNTHETASE (	3	25e+02
303	5	16.1	5	317	10	082126	3	25e+02	376	5	16.1	356	10	042625	GLUTAMINE SYNTHETASE (	3	25e+02
304	5	16.1	5	317	10	064947	3	25e+02	377	5	16.1	356	10	043785	GLUTAMINE SYNTHETASE (	3	25e+02
305	5	16.1	5	318	2	095234	3	25e+02	378	5	16.1	356	10	042951	GLUTAMINE SYNTHETASE (	3	25e+02
306	5	16.1	5	320	10	043404	3	25e+02	379	5	16.1	356	10	024335	GLUTAMINE SYNTHETASE (	3	25e+02
307	5	16.1	5	321	10	065031	3	25e+02	380	5	16.1	356	10	004999	GLUTAMINE SYNTHETASE (	3	25e+02
308	5	16.1	5	321	5	002642	3	25e+02	381	5	16.1	356	10	024334	GLUTAMINE SYNTHETASE (	3	25e+02
309	5	16.1	5	323	5	019121	3	25e+02	382	5	16.1	357	10	022094	PHOSPHOGLUCOSE ISOMERA	3	25e+02
310	5	16.1	5	326	10	049532	3	25e+02	383	5	16.1	357	10	043066	GLUTAMINE SYNTHETASE (	3	25e+02
311	5	16.1	5	326	2	046142	3	25e+02	384	5	16.1	357	10	024644	PHOSPHOGLUCOSE ISOMERA	3	25e+02
312	5	16.1	5	326	2	006638	3	25e+02	385	5	16.1	357	10	022092	PHOSPHOGLUCOSE ISOMERA	3	25e+02

386	5	16.1	357	10	022093	PHOSPHOGLUCOSE ISOMERASE	3.25e+02
387	5	16.1	357	10	022095	PHOSPHOGLUCOSE ISOMERASE	3.25e+02
388	5	16.1	358	5	016294	F32D1.5 PROTEIN.	3.25e+02
389	5	16.1	358	7	030849	MHC CLASS I ANTIGEN PR	3.25e+02
390	5	16.1	359	10	022927	APG PROTEIN ISOLOG.	3.25e+02
391	5	16.1	359	2	087161	ORF22-22 PROTEIN.	3.25e+02
392	5	16.1	359	14	072119	RNA POLYMERASE (FRAGME	3.25e+02
393	5	16.1	361	2	053294	THERMOSTABLE ALKALINE	3.25e+02
394	5	16.1	362	7	030871	MHC CLASS I B.	3.25e+02
395	5	16.1	362	10	081930	CYSSTEINE PROTEINASE.	3.25e+02
396	5	16.1	362	3	074224	SONA.	3.25e+02
397	5	16.1	362	11	061891	HISTOCOMPATIBILITY 2,	3.25e+02
398	5	16.1	362	8	020055	NATURASE (FRAGMENT).	3.25e+02
399	5	16.1	362	8	020050	NATURASE (FRAGMENT).	3.25e+02
400	5	16.1	363	5	091156	SIMILARITY TO HUMAN HE	3.25e+02
401	5	16.1	364	10	065385	FL2F1.19 PROTEIN.	3.25e+02
402	5	16.1	364	11	070509	GLYCOPROTEIN CD44S.	3.25e+02
403	5	16.1	365	7	019687	HLA-A29 VARIANT PRECUR	3.25e+02
404	5	16.1	365	11	061892	HISTOCOMPATIBILITY 2,	3.25e+02
405	5	16.1	365	1	050600	365AA LONG HYPOTHETICA	3.25e+02
406	5	16.1	365	2	027747	GLUTAMINE SYNTHETASE (	3.25e+02
407	5	16.1	367	5	031982	YOMC PROTEIN.	3.25e+02
408	5	16.1	367	5	017448	SIMILAR TO CASIN KINA	3.25e+02
409	5	16.1	367	14	072463	MAJOR OUTER CAPSID PRO	3.25e+02
410	5	16.1	367	9	064040	N-ACETYLMURAMOYL-L-ALA	3.25e+02
411	5	16.1	368	1	058192	368AA LONG HYPOTHETICA	3.25e+02
412	5	16.1	368	1	058192	368AA LONG HYPOTHETICA	3.25e+02
413	5	16.1	369	5	020804	F55C10.1 PROTEIN.	3.25e+02
414	5	16.1	370	1	029097	RIBOSE ABC TRANSPORTER	3.25e+02
415	5	16.1	370	4	075194	HUMAN PHOLACTIN-RELEAS	3.25e+02
416	5	16.1	371	2	050503	PERIPROLACTIN SOLUTE-BIN	3.25e+02
417	5	16.1	371	10	058993	371AA LONG HYPOTHETICA	3.25e+02
418	5	16.1	371	10	093771	GIBBERELLIN C-20 OXIDA	3.25e+02
419	5	16.1	371	2	047905	(ATCC 10953) MAJOR OUT	3.25e+02
420	5	16.1	372	2	059259	HISTIDINE PROTEIN KINA	3.25e+02
421	5	16.1	372	4	043344	GDF-1.	3.25e+02
422	5	16.1	373	1	026544	SENSORY TRANSDUCTION H	3.25e+02
423	5	16.1	375	2	059598	CARBAMOYL PHOSPHATE SY	3.25e+02
424	5	16.1	375	10	093403	ANTHOCYANIDIN SYNTHASE	3.25e+02
425	5	16.1	376	2	087140	ORF139-23 PROTEIN.	3.25e+02
426	5	16.1	377	2	034236	GALACTOSYL TRANSFERASE	3.25e+02
427	5	16.1	377	2	050983	CARA.	3.25e+02
428	5	16.1	379	2	072239	CFA-BETA-KETOACYLSYNTH	3.25e+02
429	5	16.1	379	4	016815	6-PHOSPHOFRUCTOKINASE,	3.25e+02
430	5	16.1	380	10	039112	GIBBERELLIN 20-OXIDASE	3.25e+02
431	5	16.1	380	2	053771	HYPOTHETICAL 41.5 KD P	3.25e+02
432	5	16.1	382	5	001300	F09C8.1 PROTEIN.	3.25e+02
433	5	16.1	382	5	044219	CEPHALOSPORINASE.	3.25e+02
434	5	16.1	383	2	025795	NA+/H+ ANTIPOITER (NAP	3.25e+02
435	5	16.1	384	2	096407	HYPOTHETICAL 42.1 KD P	3.25e+02
436	5	16.1	385	2	044330	PROTEIN KINASE (FRAGME	3.25e+02
437	5	16.1	385				







970 4 12.9 469 8 062947 6.10e+03  
971 4 12.9 473 8 062948 6.10e+03  
972 4 12.9 494 4 016352 6.10e+03  
973 4 12.9 533 6 028175 6.10e+03  
974 4 12.9 540 5 090787 6.10e+03  
975 4 12.9 554 3 074775 6.10e+03  
976 4 12.9 591 3 013607 6.10e+03  
977 4 12.9 600 5 017835 6.10e+03  
978 4 12.9 656 5 061226 6.10e+03  
979 4 12.9 695 6 077811 6.10e+03  
980 4 12.9 702 4 085050 6.10e+03  
981 4 12.9 736 3 059672 6.10e+03  
982 4 12.9 740 4 014747 6.10e+03  
983 4 12.9 740 2 05763 6.10e+03  
984 4 12.9 753 6 028959 6.10e+03  
985 4 12.9 756 2 089900 6.10e+03  
986 4 12.9 760 5 045033 6.10e+03  
987 4 12.9 764 2 031356 6.10e+03  
988 4 12.9 863 1 029837 6.10e+03  
989 4 12.9 900 5 052429 6.10e+03  
990 4 12.9 932 3 074760 6.10e+03  
991 4 12.9 1011 5 024273 6.10e+03  
992 4 12.9 1026 5 023979 6.10e+03  
993 4 12.9 1032 5 013968 6.10e+03  
994 4 12.9 1265 3 059920 6.10e+03  
995 4 12.9 1310 5 001789 6.10e+03  
996 4 12.9 1704 4 092473 6.10e+03  
997 4 12.9 1784 5 094606 6.10e+03  
998 4 12.9 1974 5 021000 6.10e+03  
999 4 12.9 3228 5 093593 6.10e+03  
1000 4 12.9 4377 4 012955 6.10e+03

## ALIGNMENTS

RESULT 1  
ID O88826 PRELIMINARY; PRT; 913 AA.  
AC O88826  
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE GOB-5 PROTEIN.  
DE GOB-5  
GN GOB-5.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RA KOMIVA T., TANIGAWA Y., HIROHASHI S.;  
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet  
cells in mice."  
RT cells in mice."  
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AB017156; D1034712;  
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 32.3%; Score 10; DB 11; Length 913;  
Best Local Similarity 100.0%; Pred. No. 4.08e-07;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 645 DNGAGADATK 654  
|||||||  
QY 1 DNGAGADATK 10

RESULT 2  
ID O05269 PRELIMINARY; PRT; 326 AA.  
AC O05269  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 35.8 KD PROTEIN.  
GN YUMD.

OS BACILLUS SUBTILIS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA OUDEGA B., KONINGSTEYN G., RAMON-DE HAAN M., RODRIGUES L.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE; 98044033.  
RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,  
RA AZEVEDO V., BERTERO M.G., BESIETTES P., BOLOFIN A., BORCHERT S.,  
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGNELL S.C., BRON S.,  
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,  
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,  
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMMERSON P.T.,  
RA ENTIAN K.D., ERRINGTON J., FABRET C., FERRARI E., FOULGER D.,  
RA FRITZ C., FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N.,  
RA GHIM S.Y., GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G.,  
RA GUISEPPI G., GUY B.J., HAGA K., HATECH J., HARWOOD C.R., HENAUT A.,  
RA HILBERT H., HOLSAPPEL S., HOSONO S., HULLO M.F., ITAYA M., JONES L.,  
RA JORIS B., KARAMATA D., KASHARA Y., KLAERR-BLANCHARD M., KLEIN C.,  
RA KOBAYASHI Y., KOETTER P., KONINGSTEIN G., KROGH S., KUMANO M.,  
RA KURIITA K., LAPIDUS A., LARDINOIS S., LAUBER J., LAZAREVIC V.,  
RA LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDIGUE C.,  
RA MEDINA N., MELLADO R.P., MIZUNO M., MOESTL D., NAKAI S., NACK M.,  
RA NOONE D., O'REILLY M., OGAWA K., OGIMARA A., OUDEGA B., PARK S.H.,  
RA PARRO V., POHL T.M., PORTELETT D., PORWOLLIK S., PRESCOTT A.M.,  
RA PRESSCAN E., PUJIC P., PURNELLE B., RAPOPORT G., REY M., REYNOLDS S.,  
RA RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAIE Y.,  
RA SATO T., SCANLAN E., SCHLEICH S., SCHROETER R., SCOPFONE E.,  
RA SEKIGUCHI J., SEKOWSKA A., SERO S.J., SERROR P., SHIN B.S., SOLDI B.,  
RA SOROKIN A., TACCONI E., TAKAGI T., TAKAHASHI H., TAKEMARU K.,  
RA TAKEUCHI M., TAKAKOSHI A., TANAKA T., TERPSTRA P., TOGNONI A.,  
RA TOSATO V., UCHIYAMA S., VANDENBOL M., VANNIER F., VASSAROTTI A.,  
RA VIARI A., WAMBUUT R., WEDLER E., WEDLER H., WEITZENEGGER T.,  
RA WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K., YASUMOTO K., YATA K.,  
RA YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E., YOSHIKAWA H., DANCHIN A.;  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
subtilis."  
RT subtilis."  
RL NATURE 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;  
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; Z93939; E311468;  
DR EMBL; Z99120; E1184292;  
DR PFAM; PF00478; IMPDH; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 326 AA; 35819 MW; 4566014F CRC32;

Query Match 22.6%; Score 7; DB 2; Length 326;  
Best Local Similarity 100.0%; Pred. No. 2.09e-01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 160 AGADATK 166  
|||||||  
QY 4 AGADATK 10

RESULT 3  
ID O25525 PRELIMINARY; PRT; 327 AA.  
AC O25525  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE GMP REDUCTASE (GUAC).  
GN HP0854.  
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).  
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;

```

OC HELICOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RX MEDLINE: 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G., B.A.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori[published erratum appears in Nature 1997 Sep
RT 25:389(6649):412].";
RL NATURE 388:539-547(1997).
DR EMBL: AE000596; G2313987; -.
DR TIGR: HP0854; -.
DR PFAM: PF00478; IMPDH: 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 327 AA; 36038 MW; FAD410A4 CRC32;

Query Match 22.6%; Score 7; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.09e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 161 AGADATK 167
Qy 4 AGADATK 10

RESULT 4
ID Q36097 PRELIMINARY; PRT: 483 AA.
AC Q36097;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DI CYTOCHROME OXIDASE SUBUNIT I (FRAGMENT).
COI.
GN THEILERIA PARVA.
OS THEILERIA PARVA.
OG MITOCHONDRION.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; PIROPLASMDIA; THEILERIIDAE;
OC THEILERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUGUGA;
RX MEDLINE: 94155854.
RA KAIRO A., FAIRLAMB A., GOBRIGHT E., NENE V.;
RT "A 7.1 kb linear DNA molecule of Theileria parva has scrambled rDNA
RT sequences and open reading frames for mitochondrially encoded
RT proteins.";
RL EMBL J. 13:898-905(1994).
DR EMBL: 223263; G437863; -.
DR PFAM: PF00115; COX1; 2.
KW MITOCHONDRION.
FT NON_TER 1
SQ SEQUENCE 483 AA; 54008 MW; 4D152FA4 CRC32;

Query Match 22.6%; Score 7; DB 8; Length 483;
Best Local Similarity 100.0%; Pred. No. 2.09e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 308 YSRYFTT 314
Qy 15 YSRYFTT 21

RESULT 5
ID O88860 PRELIMINARY; PRT: 901 AA.
AC O88860;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)

```

```

DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CHLORIDE CHANNEL CACCC.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RA ROMIO L., MUSANTE L., CINTI R., MORAN O., SERI M., GALIETTA L.J.V.;
RT "Characterization of a murine gene homologous to the bovine Cacc
RT chloride channel.";
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF052746; G3560547; -.
SQ SEQUENCE 901 AA; 100039 MW; E52BF02E CRC32;

Query Match 22.6%; Score 7; DB 11; Length 901;
Best Local Similarity 100.0%; Pred. No. 2.09e-01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 649 DNGAGAD 655
Qy 1 DNGAGAD 7

RESULT 6
ID Q12019 PRELIMINARY; PRT: 4910 AA.
AC Q12019;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE SIMILARITY NEAR N-TERMINUS TO S. CEREVISIAE DYEIN HEAVY CHAIN.
GN L8004.13.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA GEISEL C.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU 2.,
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HALLSWORTH K., HAWKINS J., HILLIER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,
RA TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C (AB972);
RA WATERSTON R.;
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RA MESSENGUY F., DUBOIS E., VIERENDELS F., SCHERENS B.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE OF 1-116 FROM N.A.
RA VERHASSELT P., VOET M., VOLCKAERT G.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [6]
RP SEQUENCE FROM N.A.
RA MIP5;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U53876; G1256854; -.
DR EMBL: 273278; E245804; -.
SQ SEQUENCE 4910 AA; 559302 MW; 22291581 CRC32;

Query Match 22.6%; Score 7; DB 3; Length 4910;

```

Best Local Similarity 100.0%; Pred. No. 2.09e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4405 GAGADAT 4411  
| | | | |  
QY 3 GAGADAT 9

RESULT 7  
ID Q39606 PRELIMINARY; PRT; 22 AA.  
AC Q39606;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE NIT-1 GENE ENCODING NITRATE REDUCTASE, PARTIAL EXON (FRAGMENT).  
OS CHLAMYDOMONAS REINHARDTII.  
OC EUKARYOTA; VIRIDIPLANTAE; CHLOROPHYTA; CHLOROPHYCEAE; VOLVOCALES;  
OC CHLAMYDOMONADACEAE; CHLAMYDOMONAS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 89367266.  
RA FERNANDEZ E., SCHNELL R., RANUM L.P.W., HUSSEY S.C., SILFLOW C.D.,  
RA LEFEBVRE P.A.;  
RT "Isolation and characterization of the nitrate reductase structural  
gene of Chlamydomonas reinhardtii."  
RL PROC. NATL. ACAD. SCI. U.S.A. 86:6449-6453(1989).  
DR EMBL: M26074; G553046; -.  
FT NON\_TER 1  
SQ SEQUENCE 22 AA; 2133 MW; 48CF360D CRC32;

Query Match 19.4%; Score 6; DB 10; Length 22;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 AGADAT 16  
| | | | |  
QY 4 AGADAT 9

RESULT 8  
ID Q46404 PRELIMINARY; PRT; 97 AA.  
AC Q46404;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE CORRESPONDS TO A 97 AMINO ACID LONG POLYPEPTIDE.  
GN LTUB.  
OS CHLAMYDIA TRACHOMATIS.  
OC BACTERIA; CHLAMYDIALES; CHLAMYDIACEAE; CHLAMYDIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LGV434;  
RX MEDLINE: 95362647.  
RA FAHR M.J., DOUGLAS A.L., XIA W., HATCH T.P.;  
RT "Characterization of late gene promoters of Chlamydia trachomatis."  
RL J. BACTERIOL. 177:4252-4260(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-D/UW-3/CX;  
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,  
RA DAVIS R.W.;  
RT "Genome sequence of an Obligate Intracellular Pathogen of Humans:  
Chlamydia trachomatis."  
RL SCIENCE 0:0-0(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-D/UW-3/CX;  
RA STEPHENS R.S., KALMAN S., LAMMEL C.J., FAN J., MARATHE R., ARAVIND L.,  
RA MITCHELL W.P., OLINGER L., TATISOV R.L., ZHAO Q., KOONIN E.V.,  
RA DAVIS R.W.;  
RT SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: L40838; G886220; -.

DR EMBL: AE001282; G3328475; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 97 AA; 11323 MW; 6A7B9AFA CRC32;

Query Match 19.4%; Score 6; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 91 GRYSVK 96  
| | | | |  
QY 26 GRYSVK 31

RESULT 9  
ID Q63144 PRELIMINARY; PRT; 104 AA.  
AC Q63144;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE OUTER MEMBRANE CYTOCHROME B (FRAGMENT).  
OS RATTUS NORVEGICUS (RAT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;  
OC SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA BORGESE N.;  
RL SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE: 95377460.  
RA DE SILVESTRIS M., D'ARRIGO A., BORGESE N.;  
RT "The targeting information of the mitochondrial outer membrane  
isoform of cytochrome b5 is contained within the carboxyl-terminal  
region."  
RL FEBS LETT. 370:69-74(1995).  
DR EMBL: X96392; E225547; -.  
DR PFAM: PF00173; heme\_1; 1.  
FT NON\_TER 1  
SQ SEQUENCE 104 AA; 11672 MW; 886FAB18 CRC32;

Query Match 19.4%; Score 6; DB 11; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 AGADAT 29  
| | | | |  
QY 4 AGADAT 9

RESULT 10  
ID Q02430 PRELIMINARY; PRT; 158 AA.  
AC Q02430;  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE CYTOSOLIC CU-ZN SUPEROXIDE DISMUTASE.  
OS DIROFILARIA IMMITIS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; SPIRURIA; SPIRURIDA;  
OC FILARIOIDEA; ONCHOCERCIDAE; DIROFILARIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA JAMES E.R., MCLEAN D.C.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: AF004949; G2209364; -.  
DR PFAM: PF00080; sodcu; 1.  
SQ SEQUENCE 158 AA; 16370 MW; 715F8D76 CRC32;

Query Match 19.4%; Score 6; DB 5; Length 158;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

Db 89 AGADAT 94
QY 4 AGADAT 9

RESULT 11
ID Q08004 PRELIMINARY; PRT: 166 AA.
AC Q08004;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE CHROMOSOME XII READING FRAME ORF YLE074C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMICETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA POHL T.M.;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Z73246; E245793; -.
SQ SEQUENCE 166 AA; 18517 MW; 5A302296 CRC32;

Query Match 19.4%; Score 6; DB 3; Length 166;
Best Local Similarity 100.0%; Pred. No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 GYSVK 7
QY 26 GYSVK 31

RESULT 12
ID Q47877 PRELIMINARY; PRT: 171 AA.
AC Q47877;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NADH DEHYDROGENASE SUBUNIT 6.
OS ALLIGATOR MISSISSIPPIENSIS (AMERICAN ALLIGATOR).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; CROCODILIA;
OC ALLIGATORIDAE; ALLIGATOR.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 98066357.
RA JANKE A., ARNASON U.;
RT "The complete mitochondrial genome of Alligator mississippiensis and
the separation between recent archosauria (birds and crocodiles).";
RL MOL. BIOL. EVOL. 14:1266-1272(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA JANKE A.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Y13113; E1253677; -.
KW MITOCHONDRION.
SQ SEQUENCE 171 AA; 19004 MW; 17636561 CRC32;

Query Match 19.4%; Score 6; DB 8; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 100 GYSRY 105
QY 13 GYSRY 18

RESULT 13

```

```

ID Q49286 PRELIMINARY; PRT: 183 AA.
AC Q49286;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE MGPA (FRAGMENT).
OS MYCOPLASMA GENITALIUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G-37;
RA PETERSON S.N., HU P., BOIT K.F., HUTCHISON C.A.;
RL J. BACTERIOL. 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-G-37;
RA PETERSON S.N.;
RL THESIS (1992), MICROBIOLOGY AND IMMUNOLOGY,
UNIVERSITY OF NORTH CAROLINA MEDICAL SCHOOL.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-G-37;
RA PETERSON S.N.;
RL SUBMITTED (SEP-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U02159; G406375; -.
FT NON_TER 1 1
SQ SEQUENCE 183 AA; 19799 MW; 8B224D2C CRC32;

Query Match 19.4%; Score 6; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.00e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 113 TTYDTN 118
QY 20 TTYDTN 25

RESULT 14
ID O86545 PRELIMINARY; PRT: 186 AA.
AC O86545;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 19.9 KD PROTEIN.
GN SC1F2.08C.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC ACTINOMYCETALES; STREPTOMYCETACEAE; STREPTOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEEGER K.J., HARRIS D.;
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE; 97000351.
RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. MICROBIOL. 21:77-96(1996).
DR EMBL: AL031350; E1316900; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 186 AA; 19874 MW; E4B54C0F CRC32;

Query Match 19.4%; Score 6; DB 2; Length 186;

```

Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 175 DDGVYS 180  
QY 11 DDGVYS 16

RESULT 15  
ID O65731 PRELIMINARY; PRT; 197 AA.  
AC O65731;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
DE 40S RIBOSOMAL PROTEIN S5 (FRAGMENT).  
OS CICER ARIETINUM (CHICKPEA) (GARBANZO).  
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC FABALES; FABACEAE; PAPILIONOIDEAE; CICER.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. CASTELLANA; TISSUE=ETIOLATED EPICOTYLS;  
RA DOPICO B., ESTEBAN R., LABRADOR E.;  
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AJ005346; E1286357;  
DR PROSITE; PS00052; RIBOSOMAL\_S7; 1.  
KW RIBOSOMAL PROTEIN.  
FT NON\_TER 1  
SQ SEQUENCE 197 AA; 22018 MW; 72B474EA CRC32;

Query Match 19.4%; Score 6; DB 10; Length 197;  
Best Local Similarity 100.0%; Pred. No. 1.00e+01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 47 GRYSVK 52  
QY 26 GRYSVK 31

Search completed: Sat Aug 28 15:02:36 1999  
Job time : 50 secs.



90	4	25.0	86	1	P93140	Synthetic TAT protein	8.13e+02	163	4	25.0	120	33	W65109	Equine chorionic gona	8.13e+02
91	4	25.0	86	24	W26443	HIV tat protein.	8.13e+02	164	4	25.0	120	2	R11984	Anti-placental alkali	8.13e+02
92	4	25.0	86	35	W76148	HIV Type I TAT protei	8.13e+02	165	4	25.0	120	1	R06278	Swine enzootic pneumo	8.13e+02
93	4	25.0	86	14	R78742	Wild type TAT HIV pro	8.13e+02	166	4	25.0	120	28	W33772	Equine chorionic gona	8.13e+02
94	4	25.0	87	3	R15192	hCG alpha subunit C-t	8.13e+02	167	4	25.0	120	3	R10039	Ovine FSH alpha subun	8.13e+02
95	4	25.0	87	2	R08154	Polyptide malarial	8.13e+02	168	4	25.0	120	28	W38115	Equine chorionic gona	8.13e+02
96	4	25.0	87	3	R13379	TAT protein.	8.13e+02	169	4	25.0	120	5	R05776	Equine alpha subunit	8.13e+02
97	4	25.0	88	19	R95256	Wild-type viral US12.	8.13e+02	170	4	25.0	122	5	R24399	Prod. of the S gene o	8.13e+02
98	4	25.0	89	3	P50638	Salmon gonadotrophin.	8.13e+02	171	4	25.0	122	22	W20110	H. pylori transmembra	8.13e+02
99	4	25.0	90	2	P81887	Sequence of human pro	8.13e+02	172	4	25.0	122	22	W24587	H. pylori transmembra	8.13e+02
100	4	25.0	90	31	W28169	Amino acid sequence o	8.13e+02	173	4	25.0	125	5	R45112	Prod. of UCD-2 of FIP	8.13e+02
101	4	25.0	91	18	R91701	AcanaPC2.	8.13e+02	174	4	25.0	127	6	R30953	Rabbit whey acidic pr	8.13e+02
102	4	25.0	91	1	P91030	Human H400 polypeptid	8.13e+02	175	4	25.0	130	1	R06282	Swine enzootic pneumo	8.13e+02
103	4	25.0	92	39	W95519	Human chorionic gonad	8.13e+02	176	4	25.0	132	1	R06277	Swine enzootic pneumo	8.13e+02
104	4	25.0	92	39	W82717	Human Act-2 protein.	8.13e+02	177	4	25.0	133	38	W85470	ATG-1120 (allograft i	8.13e+02
105	4	25.0	92	3	R15181	hCG/bCG alpha subunit	8.13e+02	178	4	25.0	138	24	W27248	Pseudomonas fluoresce	8.13e+02
106	4	25.0	92	4	R27112	Human MIP-1 alpha	8.13e+02	179	4	25.0	138	23	W08944	Heavy chain variable	8.13e+02
107	4	25.0	92	13	R70798	MIP-1-beta.	8.13e+02	180	4	25.0	138	23	W08942	Heavy chain variable	8.13e+02
108	4	25.0	92	3	R15194	hCG alpha subunit mut	8.13e+02	181	4	25.0	142	8	R47223	Charcot-Leyden Crysta	8.13e+02
109	4	25.0	92	1	R04222	PAT744 gene product e	8.13e+02	182	4	25.0	147	38	W85471	ATG-750/RC-9 (allogra	8.13e+02
110	4	25.0	92	3	R15193	hCG alpha subunit mut	8.13e+02	183	4	25.0	147	38	W85469	ATG-750/RC-9 (allogra	8.13e+02
111	4	25.0	92	1	R04167	CDNA encoding nematod	8.13e+02	184	4	25.0	147	23	W85314	Human RC-9 implicated	8.13e+02
112	4	25.0	92	7	R36770	MIP-1beta.	8.13e+02	185	4	25.0	150	13	R64904	Prorelaxin fragment	8.13e+02
113	4	25.0	92	1	R04220	Act-2 clone gene prod	8.13e+02	186	4	25.0	152	34	W77504	Staphylococcus aureus	8.13e+02
114	4	25.0	92	1	R05920	Act-2 gene product.	8.13e+02	187	4	25.0	160	33	W65071	E. coli GroEL protein	8.13e+02
115	4	25.0	92	3	R15180	hCG/bCG alpha subunit	8.13e+02	188	4	25.0	160	32	W61375	GroEL apical domain p	8.13e+02
116	4	25.0	92	35	W76225	Human chemokine MIP-1	8.13e+02	189	4	25.0	162	2	R07987	H2 prorelaxin gene en	8.13e+02
117	4	25.0	92	3	R15179	hCG/bCG alpha subunit	8.13e+02	190	4	25.0	164	2	R07988	H2 prorelaxin deduced	8.13e+02
118	4	25.0	92	1	P92146	Beta subunit of urina	8.13e+02	191	4	25.0	167	31	W28017	Staphylococcus aureus	8.13e+02
119	4	25.0	93	1	R05903	PAT 744 gene product.	8.13e+02	192	4	25.0	171	37	W26666	Canine herpes virus p	8.13e+02
120	4	25.0	95	32	W58590	Human protein AK609.1	8.13e+02	193	4	25.0	173	33	W56766	Homo sapiens LexA-PS-	8.13e+02
121	4	25.0	95	34	W77613	Deoxyribose-phosphate	8.13e+02	194	4	25.0	173	31	W28351	Staphylococcus aureus	8.13e+02
122	4	25.0	96	3	R15185	hCG/bCG alpha subunit	8.13e+02	195	4	25.0	173	22	W20487	H. pylori cytoplasmic	8.13e+02
123	4	25.0	96	3	R15184	hCG/bCG alpha subunit	8.13e+02	196	4	25.0	175	34	W62748	Streptococcus pneumon	8.13e+02
124	4	25.0	96	3	R15191	hCG/bCG alpha subunit	8.13e+02	197	4	25.0	176	13	R19433	Grb3-3 protein.	8.13e+02
125	4	25.0	96	3	R15182	hCG/bCG alpha subunit	8.13e+02	198	4	25.0	179	35	W70272	Brugia malayi transgl	8.13e+02
126	4	25.0	96	3	R15187	hCG/bCG alpha subunit	8.13e+02	199	4	25.0	184	20	W07187	B. garinii IP90 decor	8.13e+02
127	4	25.0	96	3	R15186	hCG/bCG alpha subunit	8.13e+02	200	4	25.0	185	29	W42048	Human homologue of PA	8.13e+02
128	4	25.0	96	3	R15183	hCG/bCG alpha subunit	8.13e+02	201	4	25.0	185	4	P40155	Sequence of human pre	8.13e+02
129	4	25.0	96	3	R15188	hCG/bCG alpha subunit	8.13e+02	202	4	25.0	185	4	P40154	Sequence of human pre	8.13e+02
130	4	25.0	96	3	R15189	hCG/bCG alpha subunit	8.13e+02	203	4	25.0	185	1	P94622	Amino acid sequence o	8.13e+02
131	4	25.0	96	3	R15195	hCG alpha subunit mut	8.13e+02	204	4	25.0	185	35	W69754	Ribosome recycling fa	8.13e+02
132	4	25.0	100	19	W03613	Rat non-isopeptide se	8.13e+02	205	4	25.0	185	3	R14036	Ribosome releasing fa	8.13e+02
133	4	25.0	100	3	P60256	Interferon-pseudo-ome	8.13e+02	206	4	25.0	185	4	P40108	Sequence of human pre	8.13e+02
134	4	25.0	100	2	R26668	Alpha peptide of acid	8.13e+02	207	4	25.0	185	29	W44793	B. garinii strain IP9	8.13e+02
135	4	25.0	101	3	P60357	Sequence of human leu	8.13e+02	208	4	25.0	186	33	W65089	E. coli GroEL N-termi	8.13e+02
136	4	25.0	101	3	R12259	Interferon-pseudo-ome	8.13e+02	209	4	25.0	186	36	W80402	A secreted protein en	8.13e+02
137	4	25.0	101	3	R12259	HIV-1 strain OYI TAT	8.13e+02	210	4	25.0	186	34	W77774	Staphylococcus aureus	8.13e+02
138	4	25.0	105	18	R94889	CD31 fragment (domain	8.13e+02	211	4	25.0	187	38	W89878	Antigen 1 from cluste	8.13e+02
139	4	25.0	106	21	W03989	SH2 domain from human	8.13e+02	212	4	25.0	187	38	W89817	Protein encoded by cl	8.13e+02
140	4	25.0	107	31	W59127	Human pJG4-5 library	8.13e+02	213	4	25.0	189	35	W79163	Rat calcium channel s	8.13e+02
141	4	25.0	108	35	W76214	Human occludin protei	8.13e+02	214	4	25.0	189	35	W79162	Human calcium channel	8.13e+02
142	4	25.0	113	27	W36449	Human chemokine alpha	8.13e+02	215	4	25.0	190	5	R24394	FIVP fusion protein f	8.13e+02
143	4	25.0	114	31	W46286	Human granulocyte che	8.13e+02	216	4	25.0	191	28	W30380	Candida Cacd42 prote	8.13e+02
144	4	25.0	114	2	P94149	Fish gonadotropin (GT	8.13e+02	217	4	25.0	191	28	W33897	Candida Cacd42 prote	8.13e+02
145	4	25.0	114	8	R39777	Active site peptide f	8.13e+02	218	4	25.0	191	33	W48897	Candida albicans Cacd	8.13e+02
146	4	25.0	114	10	R54672	Fish gonadotropic hor	8.13e+02	219	4	25.0	192	27	W27284	Enteropathogenic E. c	8.13e+02
147	4	25.0	114	10	R53532	Partial putative tran	8.13e+02	220	4	25.0	192	35	W68475	HIV-1 strain YBF30 vi	8.13e+02
148	4	25.0	116	4	R24805	HuRSV19VH (NIK/YCAR)	8.13e+02	221	4	25.0	194	20	W07185	B. burgdorferi HB-19	8.13e+02
149	4	25.0	116	4	R24807	RSV19 VH.	8.13e+02	222	4	25.0	194	29	W44799	B. burgdorferi strain	8.13e+02
150	4	25.0	116	2	R10095	Engineered human alph	8.13e+02	223	4	25.0	195	29	W44796	B. burgdorferi strain	8.13e+02
151	4	25.0	116	19	R99417	Dimeric glycoprotein	8.13e+02	224	4	25.0	195	29	W44801	B. burgdorferi strain	8.13e+02
152	4	25.0	116	4	R24803	HuRSV19VH (TFS/YCAR)	8.13e+02	225	4	25.0	197	21	W10221	Haemophilus influenza	8.13e+02
153	4	25.0	116	8	R42806	PhuRSV19VHFNHS.	8.13e+02	226	4	25.0	200	9	R47245	Fragment of 101 kd pr	8.13e+02
154	4	25.0	116	8	R42805	PhuRSV19VH.	8.13e+02	227	4	25.0	206	36	W79198	Mouse SNAP-25 polypep	8.13e+02
155	4	25.0	116	8	R42802	RSV19 heavy chain var	8.13e+02	228	4	25.0	206	28	W34326	Mouse synaptosomal-as	8.13e+02
156	4	25.0	116	4	R24902	HuRSV19VH.	8.13e+02	229	4	25.0	206	27	W30103	Mouse synaptosomal-as	8.13e+02
157	4	25.0	116	27	W31665	Secreted protein huma	8.13e+02	230	4	25.0	208	16	R81402	Hepatitis GB virus la	8.13e+02
158	4	25.0	116	4	R24806	HuRSV19VHNIK (NIK/FCN	8.13e+02	231	4	25.0	210	32	W58587	Mouse protein AEG48.1	8.13e+02
159	4	25.0	117	8	R40954	Germ-line VH-chain-co	8.13e+02	232	4	25.0	210	9	R46555	fbpl, for reverse gen	8.13e+02
160	4	25.0	117	5	R25524	Tobacco Ring spot Vir	8.13e+02	233	4	25.0	213	11	R60576	House dust mite aller	8.13e+02
161	4	25.0	118	9	R47248	Fragment of 101 kd pr	8.13e+02	234	4	25.0	215	11	R60575	House dust mite aller	8.13e+02
162	4	25.0	119	10	R54671	Fish gonadotropic hor	8.13e+02	235	4	25.0	215	5	R27883	Pheromone receptor cl	8.13e+02

236	4	25.0	216 34	W39501	S. pneumoniae phosphatase	8.13e+02	309	4	25.0	238 35	W59816	Green fluorescent pro	8.13e+02
237	4	25.0	217 15	R84636	Grb2 protein.	8.13e+02	310	4	25.0	238 35	W59815	Green fluorescent pro	8.13e+02
238	4	25.0	217 21	W14004	Human GRB2.	8.13e+02	311	4	25.0	238 35	W59817	Green fluorescent pro	8.13e+02
239	4	25.0	217 16	R85918	Human GRB-2.	8.13e+02	312	4	25.0	238 35	W59813	Green fluorescent pro	8.13e+02
240	4	25.0	217 29	W42070	Growth factor recepto	8.13e+02	313	4	25.0	238 35	W59814	Green fluorescent pro	8.13e+02
241	4	25.0	217 4	R20180	Sequence of the 23 kD	8.13e+02	314	4	25.0	238 35	W76371	A. victoria green flu	8.13e+02
242	4	25.0	222 16	R86256	Single chain gonadotr	8.13e+02	315	4	25.0	238 25	W24232	Aequorea victoria gre	8.13e+02
243	4	25.0	222 16	R86268	Partially deglycosyla	8.13e+02	316	4	25.0	238 33	W65082	A. victoria green flu	8.13e+02
244	4	25.0	222 16	R86278	Single chain gonadotr	8.13e+02	317	4	25.0	238 30	W52351	Engineered green fluo	8.13e+02
245	4	25.0	232 36	W72136	HSV-2 strain 95 Cont	8.13e+02	318	4	25.0	238 35	W59819	Green fluorescent pro	8.13e+02
246	4	25.0	224 36	W80708	S. pneumoniae protein	8.13e+02	319	4	25.0	238 35	W59827	Green fluorescent pro	8.13e+02
247	4	25.0	225 15	R76943	Beta 2 toxin (amino a	8.13e+02	320	4	25.0	238 30	W05312	Engineered green fluo	8.13e+02
248	4	25.0	226 16	R86277	Single chain gonadotr	8.13e+02	321	4	25.0	238 30	W52339	Engineered green fluo	8.13e+02
249	4	25.0	226 16	R86255	Single chain gonadotr	8.13e+02	322	4	25.0	238 32	W42492	A. victoria green flu	8.13e+02
250	4	25.0	228 26	R14789	VI-Lab-Vh constructio	8.13e+02	323	4	25.0	238 25	W21752	Mutant Aequorea victo	8.13e+02
251	4	25.0	228 26	R14704	VI-Lab-Vh constructio	8.13e+02	324	4	25.0	238 32	W42498	A. victoria green flu	8.13e+02
252	4	25.0	229 16	R86262	Partially deglycosyla	8.13e+02	325	4	25.0	238 32	W42494	A. victoria green flu	8.13e+02
253	4	25.0	229 16	R86250	Single chain gonadotr	8.13e+02	326	4	25.0	238 32	W42495	A. victoria green flu	8.13e+02
254	4	25.0	229 22	W20642	H. pylori cytoplasmic	8.13e+02	327	4	25.0	238 33	W65085	A. victoria green flu	8.13e+02
255	4	25.0	230 4	R20597	GAG gene p26 protein.	8.13e+02	328	4	25.0	238 33	W65080	A. victoria green flu	8.13e+02
256	4	25.0	234 16	R86275	Single chain gonadotr	8.13e+02	329	4	25.0	238 33	W65076	A. victoria green flu	8.13e+02
257	4	25.0	234 16	R86270	Single chain gonadotr	8.13e+02	330	4	25.0	238 33	W65083	A. victoria green flu	8.13e+02
258	4	25.0	234 16	R86248	Single chain gonadotr	8.13e+02	331	4	25.0	238 30	W52338	Engineered green fluo	8.13e+02
259	4	25.0	234 16	R86276	Single chain gonadotr	8.13e+02	332	4	25.0	238 20	W05311	GFP mutant Y66W/1123V	8.13e+02
260	4	25.0	234 16	R86253	Single chain gonadotr	8.13e+02	333	4	25.0	238 20	W05308	Green fluorescent pro	8.13e+02
261	4	25.0	234 16	R86260	Partially deglycosyla	8.13e+02	334	4	25.0	238 32	W42493	A. victoria green flu	8.13e+02
262	4	25.0	234 16	R86249	Single chain gonadotr	8.13e+02	335	4	25.0	238 33	W65081	A. victoria green flu	8.13e+02
263	4	25.0	234 16	R86271	Single chain gonadotr	8.13e+02	336	4	25.0	238 33	W65086	A. victoria green flu	8.13e+02
264	4	25.0	234 16	R86254	Single chain gonadotr	8.13e+02	337	4	25.0	238 20	W05305	Green fluorescent pro	8.13e+02
265	4	25.0	234 16	R86261	Partially deglycosyla	8.13e+02	338	4	25.0	238 33	W65078	A. victoria green flu	8.13e+02
266	4	25.0	235 3	R14697	VI-Lab-Vh constructio	8.13e+02	339	4	25.0	238 35	W76110	A. victoria green flu	8.13e+02
267	4	25.0	235 3	R14695	VI-Lab-Vh constructio	8.13e+02	340	4	25.0	238 21	W13909	Variant green fluore	8.13e+02
268	4	25.0	236 31	W28089	Amino acid sequence o	8.13e+02	341	4	25.0	238 35	W76105	A. victoria green flu	8.13e+02
269	4	25.0	236 3	R14702	Vh-Lab-Vl constructio	8.13e+02	342	4	25.0	238 20	W05307	Green fluorescent pro	8.13e+02
270	4	25.0	236 3	R14701	Vh-Lab-Vl constructio	8.13e+02	343	4	25.0	238 20	W05306	Green fluorescent pro	8.13e+02
271	4	25.0	236 3	R15690	VI-Lab-Vh constructio	8.13e+02	344	4	25.0	238 30	W52347	Engineered green fluo	8.13e+02
272	4	25.0	236 3	R14700	Vh-Lab-Vl constructio	8.13e+02	345	4	25.0	238 30	W52348	Engineered green fluo	8.13e+02
273	4	25.0	236 3	R14703	Vh-Lab-Vl constructio	8.13e+02	346	4	25.0	238 30	W52346	Engineered green fluo	8.13e+02
274	4	25.0	237 16	R86252	Single chain gonadotr	8.13e+02	347	4	25.0	238 30	W52331	Engineered green fluo	8.13e+02
275	4	25.0	237 16	R86263	Partially deglycosyla	8.13e+02	348	4	25.0	238 30	W52350	Engineered green fluo	8.13e+02
276	4	25.0	237 16	R86274	Single chain gonadotr	8.13e+02	349	4	25.0	238 33	W65079	A. victoria green flu	8.13e+02
277	4	25.0	237 16	R86251	Single chain gonadotr	8.13e+02	350	4	25.0	238 35	W59811	Green fluorescent pro	8.13e+02
278	4	25.0	237 16	R86264	Partially deglycosyla	8.13e+02	351	4	25.0	238 35	W76113	A. victoria green flu	8.13e+02
279	4	25.0	238 21	W13908	Variant green fluore	8.13e+02	352	4	25.0	238 35	W76112	A. victoria green flu	8.13e+02
280	4	25.0	238 32	W42488	A. victoria green flu	8.13e+02	353	4	25.0	238 35	W76106	A. victoria green flu	8.13e+02
281	4	25.0	238 32	W42496	A. victoria green flu	8.13e+02	354	4	25.0	238 30	W52343	Engineered green fluo	8.13e+02
282	4	25.0	238 32	W42486	A. victoria green flu	8.13e+02	355	4	25.0	238 30	W52344	Engineered green fluo	8.13e+02
283	4	25.0	238 30	W52335	Engineered green fluo	8.13e+02	356	4	25.0	238 36	W72066	HSV-2 strain 95 Cont	8.13e+02
284	4	25.0	238 26	W22099	Aequorea victoria pro	8.13e+02	357	4	25.0	238 35	W59818	Green fluorescent pro	8.13e+02
285	4	25.0	238 33	W61950	Green fluorescent pro	8.13e+02	358	4	25.0	238 35	W59826	Green fluorescent pro	8.13e+02
286	4	25.0	238 30	W52314	Engineered green fluo	8.13e+02	359	4	25.0	238 35	W76372	A. victoria green flu	8.13e+02
287	4	25.0	238 26	W31295	Aequorea victoria GFP	8.13e+02	360	4	25.0	238 20	W05304	Green fluorescent pro	8.13e+02
288	4	25.0	238 26	W22097	Aequorea victoria pro	8.13e+02	361	4	25.0	238 32	W42500	A. victoria green flu	8.13e+02
289	4	25.0	238 33	W65084	A. victoria green flu	8.13e+02	362	4	25.0	238 30	W52316	Engineered green fluo	8.13e+02
290	4	25.0	238 33	W61949	Green fluorescent pro	8.13e+02	363	4	25.0	238 30	W52315	Engineered green fluo	8.13e+02
291	4	25.0	238 32	W42490	A. victoria green flu	8.13e+02	364	4	25.0	238 30	W52313	Aequorea green fluore	8.13e+02
292	4	25.0	238 32	W42487	A. victoria green flu	8.13e+02	365	4	25.0	238 30	W52317	Engineered green fluo	8.13e+02
293	4	25.0	238 26	W22096	Aequorea victoria pro	8.13e+02	366	4	25.0	238 32	W42500	A. victoria green flu	8.13e+02
294	4	25.0	238 26	W22100	Aequorea victoria pro	8.13e+02	367	4	25.0	238 32	W42497	A. victoria green flu	8.13e+02
295	4	25.0	238 26	W22101	Aequorea victoria pro	8.13e+02	368	4	25.0	238 35	W76111	A. victoria green flu	8.13e+02
296	4	25.0	238 32	W42482	A. victoria green flu	8.13e+02	369	4	25.0	238 26	W22098	Aequorea victoria pro	8.13e+02
297	4	25.0	238 32	W40479	A. victoria green flu	8.13e+02	370	4	25.0	238 30	W52345	Engineered green fluo	8.13e+02
298	4	25.0	238 32	W42499	A. victoria green flu	8.13e+02	371	4	25.0	238 30	W52341	Engineered green fluo	8.13e+02
299	4	25.0	238 20	W05309	Green fluorescent pro	8.13e+02	372	4	25.0	238 30	W52342	Engineered green fluo	8.13e+02
300	4	25.0	238 20	W05310	Green fluorescent pro	8.13e+02	373	4	25.0	239 28	W41257	Blue fluorescent prot	8.13e+02
301	4	25.0	238 25	W21751	Mutant Aequorea victo	8.13e+02	374	4	25.0	239 28	W41255	Wild type green fluor	8.13e+02
302	4	25.0	238 35	W59824	Green fluorescent pro	8.13e+02	375	4	25.0	239 28	W41256	Mutant green fluore	8.13e+02
303	4	25.0	238 35	W59820	Green fluorescent pro	8.13e+02	376	4	25.0	239 25	W21753	Mutant Aequorea victo	8.13e+02
304	4	25.0	238 35	W59822	Green fluorescent pro	8.13e+02	377	4	25.0	239 32	W42502	A. victoria green flu	8.13e+02
305	4	25.0	238 35	W59821	Green fluorescent pro	8.13e+02	378	4	25.0	240 19	W05133	Single chain antibody	8.13e+02
306	4	25.0	238 35	W59812	Green fluorescent pro	8.13e+02	379	4	25.0	240 32	W42503	A. victoria green flu	8.13e+02
307	4	25.0	238 35	W59810	Green fluorescent pro	8.13e+02	380	4	25.0	241 32	W42483	A. victoria green flu	8.13e+02
308	4	25.0	238 21	W13910	Variant green fluore	8.13e+02	381	4	25.0	241 3	R12798	OV3 light and heavy	8.13e+02

382	4	25.0	241 33	W65077	A. victoria green flu	8.13e+02	455	4	25.0	325 1	R05553	Human interferon requ	8.13e+02
383	4	25.0	241 4	P40069	Sequence encoded by t	8.13e+02	456	4	25.0	326 2	R05370	IRP-1 active protein.	8.13e+02
384	4	25.0	241 32	W42504	A. victoria green flu	8.13e+02	457	4	25.0	327 23	W21666	Rat spermatid chemore	8.13e+02
385	4	25.0	242 27	W31730	Mycobacterium tubercu	8.13e+02	458	4	25.0	328 19	W02158	Aryl beta-N-acetylglu	8.13e+02
386	4	25.0	242 35	W59809	Green fluorescent pro	8.13e+02	459	4	25.0	329 15	W79961	Partial Lcrv (V antigen	8.13e+02
387	4	25.0	245 16	R81413	Hepatitis GB virus (H	8.13e+02	460	4	25.0	329 15	W01040	Y. pestis V antigen.	8.13e+02
388	4	25.0	248 25	W05396	Human SH3P18 protein.	8.13e+02	461	4	25.0	329 15	W79962	Partial Lcrv (V antigen	8.13e+02
389	4	25.0	251 22	W20602	H. pylori cytoplasmic	8.13e+02	462	4	25.0	329 19	W01041	Y. pestis V antigen.	8.13e+02
390	4	25.0	256 39	W89780	Staphylococcus aureus	8.13e+02	463	4	25.0	329 8	R44218	Murine interferon reg	8.13e+02
391	4	25.0	256 31	W36021	Phosphate starvation-	8.13e+02	464	4	25.0	329 1	R05552	Murine interferon reg	8.13e+02
392	4	25.0	256 27	W33357	TBP(20-161)/hcg-alpha	8.13e+02	465	4	25.0	329 26	W24231	Human melanoma associ	8.13e+02
393	4	25.0	257 38	W88578	Secreted protein enco	8.13e+02	466	4	25.0	331 33	W63752	A. nidulans IPNS prot	8.13e+02
394	4	25.0	260 35	W79206	Nucleus-transfer prot	8.13e+02	467	4	25.0	331 35	W76223	Human chemokine MIP-1	8.13e+02
395	4	25.0	260 3	P61317	Serratia nuclease prot	8.13e+02	468	4	25.0	332 17	R95449	Human abortive infect	8.13e+02
396	4	25.0	261 1	P90438	Protein containing ep	8.13e+02	469	4	25.0	332 17	R88353	Mutant A-53-Q fragmen	8.13e+02
397	4	25.0	263 21	W02307	Modified Green Fluore	8.13e+02	470	4	25.0	333 25	W33108	Chicken lactic acid d	8.13e+02
398	4	25.0	263 21	W02308	Modified Green Fluore	8.13e+02	471	4	25.0	333 25	W33107	Chicken lactic acid d	8.13e+02
399	4	25.0	263 21	W02309	Modified Green Fluore	8.13e+02	472	4	25.0	334 38	W89915	Antigen 2 from cluste	8.13e+02
400	4	25.0	265 38	W89987	Expressed antigen for	8.13e+02	473	4	25.0	336 1	R06280	Swine enzootic pneumo	8.13e+02
401	4	25.0	265 16	R66259	Partially deglycosyla	8.13e+02	474	4	25.0	336 3	R12312	Partial sequence of O	8.13e+02
402	4	25.0	265 22	W0929	H. pylori surface or	8.13e+02	475	4	25.0	337 13	R71035	Human IFN-gamma acces	8.13e+02
403	4	25.0	265 16	R66247	Single chain gonadotr	8.13e+02	476	4	25.0	337 1	R06283	Swine enzootic pneumo	8.13e+02
404	4	25.0	266 24	W11896	MACH isoform alpha3.	8.13e+02	477	4	25.0	340 24	W26431	Swinepox virus HindII	8.13e+02
405	4	25.0	266 34	W69739	Human integral membra	8.13e+02	478	4	25.0	341 32	W61758	B. burgdorferi antige	8.13e+02
406	4	25.0	266 32	W63690	Human secreted protei	8.13e+02	479	4	25.0	341 6	R33280	P39-beta.	8.13e+02
407	4	25.0	266 33	W55069	Streptococcus pneumon	8.13e+02	480	4	25.0	341 30	W47430	Amino acid sequence o	8.13e+02
408	4	25.0	267 39	W67885	Human secreted protei	8.13e+02	481	4	25.0	343 21	W01551	MOCl protein.	8.13e+02
409	4	25.0	268 18	R90682	Rat caldesmon contg.	8.13e+02	482	4	25.0	344 1	P90528	B cell stimulating fa	8.13e+02
410	4	25.0	269 14	R2015	Variant lipase of Hum	8.13e+02	483	4	25.0	347 2	R11069	12D3 antigen sequence	8.13e+02
411	4	25.0	269 26	R64140	Heymann nephritis ant	8.13e+02	484	4	25.0	348 29	W45575	Pyrococcus furiosus d	8.13e+02
412	4	25.0	270 38	W88784	Polypeptide fragment	8.13e+02	485	4	25.0	348 17	R93116	Rhodopsin.	8.13e+02
413	4	25.0	270 24	W15252	Asialoglycoprotein re	8.13e+02	486	4	25.0	348 7	R38483	Rhodopsin protein.	8.13e+02
414	4	25.0	272 30	W40215	Human macrophage anti	8.13e+02	487	4	25.0	348 29	W45558	Pyrococcus furiosus d	8.13e+02
415	4	25.0	273 22	W20667	H. pylori cytoplasmic	8.13e+02	488	4	25.0	349 1	P90554	Bovine rhodopsin.	8.13e+02
416	4	25.0	274 37	W79271	Phenolic acid esteras	8.13e+02	489	4	25.0	351 19	W01052	Human umbilical vein	8.13e+02
417	4	25.0	281 1	R06279	Swine enzootic pneumo	8.13e+02	490	4	25.0	354 26	W24254	Aquifex histidinol-ph	8.13e+02
418	4	25.0	283 27	W37482	Mouse liver cancer-or	8.13e+02	491	4	25.0	354 15	R75210	Alteromonas hanelai	8.13e+02
419	4	25.0	284 38	W85102	Thyroid hormone recep	8.13e+02	492	4	25.0	355 7	R33420	Human IL-8 receptor f	8.13e+02
420	4	25.0	284 19	W02712	G-protein coupled odo	8.13e+02	493	4	25.0	356 30	W46917	Amino acid sequence o	8.13e+02
421	4	25.0	284 16	R48740	G-protein coupled odo	8.13e+02	494	4	25.0	357 32	W60226	Bacillus thuringiens	8.13e+02
422	4	25.0	285 27	R33359	TBP(20-190)/hcg-alpha	8.13e+02	495	4	25.0	357 32	W60227	Bacillus thuringiens	8.13e+02
423	4	25.0	287 24	W15246	Asialoglycoprotein re	8.13e+02	496	4	25.0	360 38	W70804	Amino acid sequence o	8.13e+02
424	4	25.0	287 10	R53544	Thyroid hormone recep	8.13e+02	497	4	25.0	360 15	R80953	Recombinant high affi	8.13e+02
425	4	25.0	288 39	W89984	Expressed antigen for	8.13e+02	498	4	25.0	360 15	R80758	Interleukin 8 recept	8.13e+02
426	4	25.0	289 38	W85092	Thyroid hormone recep	8.13e+02	499	4	25.0	360 6	R28273	Sequence in a low aff	8.13e+02
427	4	25.0	292 21	W67933	ILTV unique short reg	8.13e+02	500	4	25.0	362 21	W06787	ILRV glycoprotein g.	8.13e+02
428	4	25.0	292 19	W00641	Infectious laryngotra	8.13e+02	501	4	25.0	362 19	W00635	ILRV glycoprotein g.	8.13e+02
429	4	25.0	293 34	W71481	Helicobacter polypept	8.13e+02	502	4	25.0	362 8	R44402	Canine coronavirus 1-	8.13e+02
430	4	25.0	295 10	R53534	Thyroid hormone recep	8.13e+02	503	4	25.0	363 23	W22066	Chick fringe B (lunat	8.13e+02
431	4	25.0	296 25	W05397	Human clone 55 protei	8.13e+02	504	4	25.0	364 39	W95356	IL-1 alpha proplece-E	8.13e+02
432	4	25.0	297 17	R15599	Streptokinase from PS	8.13e+02	505	4	25.0	366 17	R95177	CONSTANS protein.	8.13e+02
433	4	25.0	299 39	W89198	Aspartate-specific cy	8.13e+02	506	4	25.0	369 22	W20779	H. pylori flagella-as	8.13e+02
434	4	25.0	303 38	W89842	Protein encoded by cl	8.13e+02	507	4	25.0	371 38	W67443	C. jejuni flagellin/E	8.13e+02
435	4	25.0	304 38	W89919	Antigen 4 from cluste	8.13e+02	508	4	25.0	372 39	W67857	Human secreted protei	8.13e+02
436	4	25.0	304 18	R94891	CD31 fragment (domain	8.13e+02	509	4	25.0	373 10	R55571	Tomato ABH2.	8.13e+02
437	4	25.0	306 27	W34526	htCP protein fragment	8.13e+02	510	4	25.0	374 39	W88206	Alcohol dehydrogenase	8.13e+02
438	4	25.0	309 38	W89960	Antigen from cluster	8.13e+02	511	4	25.0	374 39	W88205	Alcohol dehydrogenase	8.13e+02
439	4	25.0	309 16	R48735	G-protein coupled hum	8.13e+02	512	4	25.0	374 39	W88207	Alcohol dehydrogenase	8.13e+02
440	4	25.0	309 19	W02707	G-protein coupled hum	8.13e+02	513	4	25.0	374 39	W88208	Horse liver alcohol d	8.13e+02
441	4	25.0	309 22	W20702	H. pylori cytoplasmic	8.13e+02	514	4	25.0	374 39	W88201	Alcohol dehydrogenase	8.13e+02
442	4	25.0	312 21	W06933	CagI locus product 4.	8.13e+02	515	4	25.0	374 39	W88200	Alcohol dehydrogenase	8.13e+02
443	4	25.0	312 5	R26079	Putative tat-E2 fusio	8.13e+02	516	4	25.0	374 39	W88204	Alcohol dehydrogenase	8.13e+02
444	4	25.0	317 16	R48685	G-protein coupled dic	8.13e+02	517	4	25.0	374 39	W88203	Alcohol dehydrogenase	8.13e+02
445	4	25.0	317 33	W60997	Streptococcus pneumon	8.13e+02	518	4	25.0	374 39	W88209	Horse liver alcohol d	8.13e+02
446	4	25.0	317 19	W02657	G-protein coupled Dic	8.13e+02	519	4	25.0	374 39	W88202	Alcohol dehydrogenase	8.13e+02
447	4	25.0	318 38	W89997	Expressed antigen for	8.13e+02	520	4	25.0	375 38	W84085	Human membrane fusio	8.13e+02
448	4	25.0	319 3	P61318	Serratia phospholipas	8.13e+02	521	4	25.0	376 22	W04470	Rat geranylgeranyl tr	8.13e+02
449	4	25.0	319 26	W22104	Murine leptin recepto	8.13e+02	522	4	25.0	376 37	W76765	R. sphaeroides AdHI c	8.13e+02
450	4	25.0	321 35	W71292	Protein Septin-2 pep.	8.13e+02	523	4	25.0	376 31	W46287	Rhizobium meliloti DN	8.13e+02
451	4	25.0	323 34	W64461	Human secreted protei	8.13e+02	524	4	25.0	376 37	W72976	Rhodobacter sphaeroid	8.13e+02
452	4	25.0	323 17	R90570	Cysteine-protease.	8.13e+02	525	4	25.0	376 22	W20307	H. pylori secreted or	8.13e+02
453	4	25.0	325 8	R44219	Human interferon requ	8.13e+02	526	4	25.0	377 22	W04469	Human geranylgeranyl	8.13e+02
454	4	25.0	325 35	W77149	Murine interleukin-18	8.13e+02	527	4	25.0	377 31	W46288	Brucella abortus DNA	8.13e+02

528	4	25.0	379	38	W85036	Alpha-actinin actin-b	8.13e+02	601	4	25.0	468	7	R37215	IL-6 receptor.	8.13e+02
529	4	25.0	380	39	W89183	S. pneumoniae GlmU OR	8.13e+02	602	4	25.0	468	19	R98364	Interleukin-6 recepto	8.13e+02
530	4	25.0	382	39	W88491	Human liver clone HP0	8.13e+02	603	4	25.0	468	37	W71371	Human interleukin-6 r	8.13e+02
531	4	25.0	382	37	W80323	Bacillus thuringiensis	8.13e+02	604	4	25.0	468	1	P90525	B cell stimulating fa	8.13e+02
532	4	25.0	382	22	W20701	H. pylori secreted or	8.13e+02	605	4	25.0	468	4	P90284	Sequence of a recepto	8.13e+02
533	4	25.0	382	21	W06932	CagI locus product 3.	8.13e+02	606	4	25.0	469	20	W04722	Aromatic acyl transfe	8.13e+02
534	4	25.0	385	37	W80436	Feline herpesvirus pr	8.13e+02	607	4	25.0	470	4	R20255	Sequence encoded by t	8.13e+02
535	4	25.0	386	1	P90526	B cell stimulating fa	8.13e+02	608	4	25.0	470	13	R66240	ADP-glucose-pyrophos	8.13e+02
536	4	25.0	389	30	W53163	Arthrobacter sarcosin	8.13e+02	609	4	25.0	470	13	R66240	ADPGPP large subunit.	8.13e+02
537	4	25.0	389	30	W53158	Arthrobacter sarcosin	8.13e+02	610	4	25.0	471	25	W16794	Aspartase-470.	8.13e+02
538	4	25.0	389	30	W53159	Arthrobacter sarcosin	8.13e+02	611	4	25.0	474	18	R94893	CD31 fragment (domain	8.13e+02
539	4	25.0	389	10	R52658	Arthrobacter sp. FE18	8.13e+02	612	4	25.0	475	26	W24228	Human melanoma associ	8.13e+02
540	4	25.0	389	30	W53156	Arthrobacter sarcosin	8.13e+02	613	4	25.0	475	34	W68154	Human lipoprotein lip	8.13e+02
541	4	25.0	389	24	W25149	Modified Arthrobacter	8.13e+02	614	4	25.0	476	14	R70099	Lettuce infectious ye	8.13e+02
542	4	25.0	389	24	W25150	Modified Arthrobacter	8.13e+02	615	4	25.0	477	25	W22161	ApxIID protein.	8.13e+02
543	4	25.0	389	24	W25151	Modified Arthrobacter	8.13e+02	616	4	25.0	479	36	W75963	Human Caspase-8.	8.13e+02
544	4	25.0	389	30	W53164	Arthrobacter sarcosin	8.13e+02	617	4	25.0	479	28	W31523	Interleukin-1 beta co	8.13e+02
545	4	25.0	389	30	W53162	Arthrobacter sarcosin	8.13e+02	618	4	25.0	480	19	W01045	Y. pestis FI/V antige	8.13e+02
546	4	25.0	389	15	R76735	N-methyl valine oxida	8.13e+02	619	4	25.0	481	9	R47246	Recombinant protein p	8.13e+02
547	4	25.0	389	30	W53155	Arthrobacter wild-typ	8.13e+02	620	4	25.0	481	31	W59045	Human MMR1-F3 protei	8.13e+02
548	4	25.0	389	30	W53161	Arthrobacter sarcosin	8.13e+02	621	4	25.0	483	3	R12522	Factor Xa-cleavable s	8.13e+02
549	4	25.0	389	36	W71462	A modified sarcosine	8.13e+02	622	4	25.0	483	3	R12885	Factor Xa-cleavable h	8.13e+02
550	4	25.0	389	36	W71461	A modified sarcosine	8.13e+02	623	4	25.0	490	8	R41663	Paired basic amino ac	8.13e+02
551	4	25.0	389	14	R79150	Arthrobacter sp. FE18	8.13e+02	624	4	25.0	492	30	W37733	Cytochrome zp450RAI p	8.13e+02
552	4	25.0	390	20	W07397	Bacillus sarcosine ox	8.13e+02	625	4	25.0	492	21	W01560	CfBI protein.	8.13e+02
553	4	25.0	390	14	R83825	p47(phox) protein.	8.13e+02	626	4	25.0	493	1	R04883	Human prolidase.	8.13e+02
554	4	25.0	390	23	W21678	Haemophilus influenza	8.13e+02	627	4	25.0	494	27	W30711	Human ubiqlutin-speci	8.13e+02
555	4	25.0	390	1	P94663	New sarcosine oxidase	8.13e+02	628	4	25.0	494	34	W69250	Human p56-2 protein.	8.13e+02
556	4	25.0	396	15	R76550	Thermotable alkaline	8.13e+02	629	4	25.0	496	27	W27391	Mch5 protein.	8.13e+02
557	4	25.0	401	21	W11156	Calreticulin.	8.13e+02	630	4	25.0	496	32	W47527	Amino acid sequence o	8.13e+02
558	4	25.0	401	36	W22779	Human septin-2 protei	8.13e+02	631	4	25.0	496	32	W47532	Human TIE-2 ligand 2	8.13e+02
559	4	25.0	403	1	P91165	38 kd regression-asso	8.13e+02	632	4	25.0	496	13	R76075	Carbamoylphosphate-sy	8.13e+02
560	4	25.0	403	8	R40855	38kd regression assoc	8.13e+02	633	4	25.0	496	20	W01411	Human TIE-2 ligand 2.	8.13e+02
561	4	25.0	403	19	W04171	Flea calreticulin Pct	8.13e+02	634	4	25.0	499	32	W47529	Amino acid sequence o	8.13e+02
562	4	25.0	404	34	W62620	Mus musculus SOCS6 pr	8.13e+02	635	4	25.0	499	3	R12891	Streptokinase fused t	8.13e+02
563	4	25.0	406	3	R12395	Transcription activat	8.13e+02	636	4	25.0	501	25	W31880	GFP variants P4-3 and	8.13e+02
564	4	25.0	411	26	W24229	Human melanoma associ	8.13e+02	637	4	25.0	501	25	W31877	GFP variants S65C and	8.13e+02
565	4	25.0	414	39	W86143	Streptokinase (SK) pr	8.13e+02	638	4	25.0	501	19	W01044	Y. pestis FI/V antige	8.13e+02
566	4	25.0	414	39	W86144	De-immunised streptok	8.13e+02	639	4	25.0	501	25	W31879	GFP variants S65T and	8.13e+02
567	4	25.0	414	4	R20202	S. equisimilis streptok	8.13e+02	640	4	25.0	502	28	W44153	Human neuronal nicoti	8.13e+02
568	4	25.0	417	1	P92276	60 kD Ro (Ro/SSA) ant	8.13e+02	641	4	25.0	502	33	W69216	V274T variant human a	8.13e+02
569	4	25.0	419	19	W01037	Mycoplasma 46-48 kDa	8.13e+02	642	4	25.0	502	20	W09025	Neuronal nicotinic ac	8.13e+02
570	4	25.0	419	4	R21829	Sequence of surface a	8.13e+02	643	4	25.0	502	21	W12368	Neuronal alpha-bungar	8.13e+02
571	4	25.0	420	33	W55075	Streptococcus pneumon	8.13e+02	644	4	25.0	504	28	W44156	Human neuronal nicoti	8.13e+02
572	4	25.0	420	35	W71291	Protein CDC pep.	8.13e+02	645	4	25.0	504	20	W09022	Neuronal nicotinic ac	8.13e+02
573	4	25.0	422	38	W73363	Chicken PAF-AH protei	8.13e+02	646	4	25.0	507	22	T21010	H. pylori chaperone p	8.13e+02
574	4	25.0	422	25	W26503	Chicken platelet-acti	8.13e+02	647	4	25.0	509	17	R97645	Human SOX-9.	8.13e+02
575	4	25.0	429	22	W20878	H. pylori transporter	8.13e+02	648	4	25.0	509	25	W05399	Human clone 65 protei	8.13e+02
576	4	25.0	431	37	W80310	Human PACAP/VIP R-2B	8.13e+02	649	4	25.0	509	17	R97645	Mouse Sox-9 protein.	8.13e+02
577	4	25.0	433	39	W89784	Staphylococcus aureus	8.13e+02	650	4	25.0	512	28	W44283	Mouse sperm surface p	8.13e+02
578	4	25.0	434	21	W13406	Solanum melongena fla	8.13e+02	651	4	25.0	512	8	R47347	Murine PH-20.	8.13e+02
579	4	25.0	438	36	W80308	dnaA gene product of	8.13e+02	652	4	25.0	513	15	R79945	Helicobacter pylori a	8.13e+02
580	4	25.0	438	37	W59995	Human PACAP/VIP R-2 r	8.13e+02	653	4	25.0	514	25	W31876	GFP variants S65C and	8.13e+02
581	4	25.0	439	37	W59995	Cell division Divib p	8.13e+02	654	4	25.0	515	33	W55073	Streptococcus pneumon	8.13e+02
582	4	25.0	440	39	W86320	Kidney injury associa	8.13e+02	655	4	25.0	516	1	P80808	Sequence of gag prote	8.13e+02
583	4	25.0	440	3	R12889	Streptokinase.	8.13e+02	656	4	25.0	519	35	W77140	Murine interleukin-18	8.13e+02
584	4	25.0	441	4	R25394	Sequence encodes endo	8.13e+02	657	4	25.0	519	35	W70285	Oryza sativa RNA poly	8.13e+02
585	4	25.0	442	38	W85035	Green fluorescent pro	8.13e+02	658	4	25.0	520	36	W72642	Modified alkaline pho	8.13e+02
586	4	25.0	442	17	W00505	N meningitidis IM2394	8.13e+02	659	4	25.0	521	23	R13053	HIV-2 provirus-encode	8.13e+02
587	4	25.0	443	38	W89850	Antigen 1 from cluste	8.13e+02	660	4	25.0	521	13	R69973	ADP-glucose-pyrophosp	8.13e+02
588	4	25.0	448	10	R51479	Human TIE-4.	8.13e+02	661	4	25.0	521	13	R66239	ADPGPP small subunit.	8.13e+02
589	4	25.0	451	39	W89977	Protein encoded by cl	8.13e+02	662	4	25.0	521	36	R20254	Sequence encoded by t	8.13e+02
590	4	25.0	451	39	W89977	Homo sapiens PSP1-1 s	8.13e+02	663	4	25.0	522	36	W72643	Alkaline phosphatase.	8.13e+02
591	4	25.0	458	33	W56777	Homo sapiens PSP1.	8.13e+02	664	4	25.0	522	1	P80801	HIV virus gag gene.	8.13e+02
592	4	25.0	458	33	W56769	Homo sapiens PSP1-1 c	8.13e+02	665	4	25.0	522	16	R84552	Sequence encoded by o	8.13e+02
593	4	25.0	458	33	W56773	Homo sapiens PSP1-1 c	8.13e+02	666	4	25.0	522	1	P82452	Hepatitis GB virus (H	8.13e+02
594	4	25.0	458	33	W56772	Homo sapiens PSP1 con	8.13e+02	667	4	25.0	522	1	P82452	GAGRODN sequence from	8.13e+02
595	4	25.0	459	39	W89182	S. pneumoniae GlmU po	8.13e+02	668	4	25.0	523	6	W23776	Human septin-2 protei	8.13e+02
596	4	25.0	461	7	R39354	EpiP protein.	8.13e+02	669	4	25.0	526	6	R33375	Rebivacterium flavum	8.13e+02
597	4	25.0	462	35	W68397	Clostridium botulinum	8.13e+02	670	4	25.0	529	39	W73473	Human cancer-related	8.13e+02
598	4	25.0	464	24	W18895	MACH isoform alphas2.	8.13e+02	671	4	25.0	536	37	W79272	Phenolic acid esteras	8.13e+02
599	4	25.0	464	37	W80420	CD2 associated intrac	8.13e+02	672	4	25.0	537	26	W31908	Mouse receptor protei	8.13e+02
600	4	25.0	465	33	W56295	Babesia microti BMNI-	8.13e+02	673	4	25.0	539	30	W33628	Yeast transcriptional	8.13e+02

674	4	25.0	544 38	W85029	Green fluorescent pro	8.13e+02
675	4	25.0	544 38	W85028	CDK2-green fluorescen	8.13e+02
676	4	25.0	544 31	W51256	Maize proto-porphyrin	8.13e+02
677	4	25.0	544 26	W25748	Maize protoporphyrino	8.13e+02
678	4	25.0	544 16	R90298	Maize protox-2 partia	8.13e+02
679	4	25.0	544 35	W71203	Protein encoded by OR	8.13e+02
680	4	25.0	545 39	W89277	Granulocytic Ehrlich	8.13e+02
681	4	25.0	545 8	R42742	ced-4 gene product.	8.13e+02
682	4	25.0	545 16	R79579	Mouse COT-gamma subu	8.13e+02
683	4	25.0	548 12	R67382	E. coli GroEL gene pr	8.13e+02
684	4	25.0	548 12	R61378	GroEL N-terminal regi	8.13e+02
685	4	25.0	548 33	W63070	E. coli GroEL protein	8.13e+02
686	4	25.0	552 1	R07562	Polypeptide encoded b	8.13e+02
687	4	25.0	553 25	W26495	CD2 associated intrac	8.13e+02
688	4	25.0	553 36	W71869	Amino acid encoded by	8.13e+02
689	4	25.0	553 37	W80419	CD2 associated intrac	8.13e+02
690	4	25.0	553 42	W25115	CD2-associated intrac	8.13e+02
691	4	25.0	553 35	W69385	Prostate tumour speci	8.13e+02
692	4	25.0	554 31	W48661	RG fusion protein seq	8.13e+02
693	4	25.0	558 31	W48662	GR fusion protein seq	8.13e+02
694	4	25.0	560 22	W18310	Human complement C9.	8.13e+02
695	4	25.0	560 3	R15051	Hybrid human insulin-	8.13e+02
696	4	25.0	561 21	W14503	Saccharomyces cerevis	8.13e+02
697	4	25.0	562 23	W12096	Adenovirus E1A associ	8.13e+02
698	4	25.0	565 14	R76699	ShET2 enterotoxin enc	8.13e+02
699	4	25.0	565 14	R76700	ETET enterotoxin enco	8.13e+02
700	4	25.0	571 18	R94894	CD31 fragment (domai	8.13e+02
701	4	25.0	572 33	W61161	Human squalene epoxid	8.13e+02
702	4	25.0	573 17	R75741	B31 outer surface pro	8.13e+02
703	4	25.0	575 1	R04714	Amino acid sequence o	8.13e+02
704	4	25.0	578 16	R18844	Human afamin mature p	8.13e+02
705	4	25.0	579 7	R34446	N.meningitidis 2394.T	8.13e+02
706	4	25.0	580 18	R94583	C. pneumoniae polypep	8.13e+02
707	4	25.0	580 20	W01742	C. pneumoniae 73 kDa	8.13e+02
708	4	25.0	580 20	W11707	C. pneumoniae 73K ant	8.13e+02
709	4	25.0	581 26	W22103	Murine leptin recepto	8.13e+02
710	4	25.0	584 6	R30811	Feline parvovirus cap	8.13e+02
711	4	25.0	584 35	R78789	Parvovirus typ	8.13e+02
712	4	25.0	585 27	W31271	Human frizzled-5 prot	8.13e+02
713	4	25.0	586 9	R45579	Sequence of bovine ep	8.13e+02
714	4	25.0	588 9	R47257	Pre-pro-DPP.	8.13e+02
715	4	25.0	592 38	W07097	Human Interleukin-6R-	8.13e+02
716	4	25.0	595 38	W85021	mPKAc-green fluoresce	8.13e+02
717	4	25.0	597 2	R12384	Bovine pancreatic cho	8.13e+02
718	4	25.0	599 9	R48219	N.meningitidis IM2394	8.13e+02
719	4	25.0	599 17	R88644	Neisseria meningitidi	8.13e+02
720	4	25.0	599 14	R73655	Human afamin.	8.13e+02
721	4	25.0	599 16	R81845	Human afamin complete	8.13e+02
722	4	25.0	599 26	W14620	HTR Tbp2 protein from	8.13e+02
723	4	25.0	600 5	R09419	Hsp70 antigen of Myco	8.13e+02
724	4	25.0	600 5	R03922	M.hypopneumoniae HSP	8.13e+02
725	4	25.0	604 38	W85016	Erk2-green fluorescen	8.13e+02
726	4	25.0	605 38	W85007	Erk2-green fluorescen	8.13e+02
727	4	25.0	605 34	W62838	Glycine max antimicro	8.13e+02
728	4	25.0	606 38	W85019	p38-green fluorescen	8.13e+02
729	4	25.0	606 18	R99673	Receptor for hyaluron	8.13e+02
730	4	25.0	607 38	W85010	p38-green fluorescen	8.13e+02
731	4	25.0	608 5	R25246	Chicken Growth Hormon	8.13e+02
732	4	25.0	617 24	W27247	Pseudomonas fluoresce	8.13e+02
733	4	25.0	619 15	R90697	Borrelia afzelii 66 k	8.13e+02
734	4	25.0	621 27	W36451	Human tyrosine phosph	8.13e+02
735	4	25.0	621 17	R90698	Borrelia garinii 66 k	8.13e+02
736	4	25.0	623 38	W88707	Secreted protein enco	8.13e+02
737	4	25.0	624 38	W85015	Erk1-green fluorescen	8.13e+02
738	4	25.0	630 38	W85018	Erk1-green fluorescen	8.13e+02
739	4	25.0	630 28	W09969	Amino acid sequence o	8.13e+02
740	4	25.0	630 30	W54127	H. influenzae strain	8.13e+02
741	4	25.0	630 27	W39166	Mouse RHAMM protein.	8.13e+02
742	4	25.0	631 38	W85006	Erk1-green fluorescen	8.13e+02
743	4	25.0	631 18	R77896	Bacterial transferrin	8.13e+02
744	4	25.0	631 28	W08970	Amino acid sequence o	8.13e+02
745	4	25.0	631 30	W54128	H. influenzae strain	8.13e+02
746	4	25.0	631 18	R99675	RHAMM 1-2a isoform.	8.13e+02

747	4	25.0	633 38	W85009	Jnk1-green fluorescen	8.13e+02
748	4	25.0	634 29	W26785	Nudarelia beta-like	8.13e+02
749	4	25.0	634 29	W34537	Nudarelia beta virus	8.13e+02
750	4	25.0	635 38	W85034	Green fluorescent pro	8.13e+02
751	4	25.0	635 14	R75940	Human myeloproliferat	8.13e+02
752	4	25.0	635 4	R23970	MPLV env protein with	8.13e+02
753	4	25.0	642 34	W62596	Human nuclear protein	8.13e+02
754	4	25.0	642 36	W71647	Fluorescent calmoduli	8.13e+02
755	4	25.0	642 37	W80950	Amino acid sequence o	8.13e+02
756	4	25.0	642 36	W71645	Fluorescent calmoduli	8.13e+02
757	4	25.0	649 31	W53827	Pseudomonas XcpQ secr	8.13e+02
758	4	25.0	652 15	R88124	Tobacco mosaic virus	8.13e+02
759	4	25.0	652 36	W71646	Fluorescent calmoduli	8.13e+02
760	4	25.0	652 3	R14047	B.thuringiensis CryII	8.13e+02
761	4	25.0	655 30	W48265	Sf9 alpha-mannosidase	8.13e+02
762	4	25.0	655 30	W46900	Sf9 alpha-mannosidase	8.13e+02
763	4	25.0	656 36	W71648	Fluorescent calmoduli	8.13e+02
764	4	25.0	659 29	W37724	CD2 associated intrac	8.13e+02
765	4	25.0	663 3	R13139	B.burgdorferi strain	8.13e+02
766	4	25.0	663 18	R88732	S.aureus topoisomeras	8.13e+02
767	4	25.0	672 39	W73536	MEKK1 protein.	8.13e+02
768	4	25.0	672 15	R77544	MEKK1 protein.	8.13e+02
769	4	25.0	672 12	R66029	Mammalian MEK kinase	8.13e+02
770	4	25.0	678 36	W72244	HSV-2 strain SB5 Cont	8.13e+02
771	4	25.0	692 8	R43581	Plant NADPH cytochrom	8.13e+02
772	4	25.0	694 1	R04107	DNA-binding protein G	8.13e+02
773	4	25.0	700 17	R75731	B. burgdorferi strain	8.13e+02
774	4	25.0	700 1	R06366	Human alpha-1 collage	8.13e+02
775	4	25.0	707 2	R07077	92-kDa human type IV	8.13e+02
776	4	25.0	708 33	W69844	Amino acid sequence o	8.13e+02
777	4	25.0	712 8	R43582	Plant NADPH cytochrom	8.13e+02
778	4	25.0	713 17	R85878	WD-40 domain-contg. T	8.13e+02
779	4	25.0	714 9	R47244	Fragment of 101 kD pr	8.13e+02
780	4	25.0	716 6	R30730	B. burgdorferi 79 kD	8.13e+02
781	4	25.0	718 38	W85024	Erk2-green fluorescen	8.13e+02
782	4	25.0	718 17	R85865	WD-40 domain-contg. D	8.13e+02
783	4	25.0	718 36	W72245	HSV-2 strain SB5 Cont	8.13e+02
784	4	25.0	719 38	W85012	Smad2-green fluoresce	8.13e+02
785	4	25.0	719 10	R52955	Drosophila E(spl)m9/1	8.13e+02
786	4	25.0	719 10	R51481	Drosophila E(spl)m9/1	8.13e+02
787	4	25.0	725 20	W01460	Hib Hxuc protein.	8.13e+02
788	4	25.0	726 38	W85022	PKB-green fluorescen	8.13e+02
789	4	25.0	727 38	W85041	Green fluorescent pro	8.13e+02
790	4	25.0	728 36	W72246	HSV-2 strain SB5 Cont	8.13e+02
791	4	25.0	738 3	R13251	PECAM-1.	8.13e+02
792	4	25.0	738 23	W14802	PECAM-1.	8.13e+02
793	4	25.0	740 9	R47174	Predicted sequence of	8.13e+02
794	4	25.0	743 10	R31110	Human TLE-2.	8.13e+02
795	4	25.0	743 10	R31477	Human TLE-2.	8.13e+02
796	4	25.0	745 2	R11809	RNA dependant RNA pol	8.13e+02
797	4	25.0	745 1	P90615	Polypeptide with isoa	8.13e+02
798	4	25.0	747 3	R12894	Met-core streptokinas	8.13e+02
799	4	25.0	748 5	R24396	Prod. of the S gene o	8.13e+02
800	4	25.0	748 5	R24398	Prod. of the S gene o	8.13e+02
801	4	25.0	748 5	R24513	Consensus sequence en	8.13e+02
802	4	25.0	749 9	R42465	Feline infectious per	8.13e+02
803	4	25.0	749 9	R42466	Feline infectious per	8.13e+02
804	4	25.0	750 1	P94419	Primary transcript of	8.13e+02
805	4	25.0	751 4	R23582	Branching enzyme.	8.13e+02
806	4	25.0	752 27	W34178	Human transcription r	8.13e+02
807	4	25.0	753 39	W83927	Human T85 protein.	8.13e+02
808	4	25.0	758 16	R86998	Enterohaemorrhagic E.	8.13e+02
809	4	25.0	765 38	W89914	Antigen from cluster	8.13e+02
810	4	25.0	770 10	R51109	Human TLE-1 protein.	8.13e+02
811	4	25.0	770 10	R51476	Human TLE-1.	8.13e+02
812	4	25.0	771 36	W72247	HSV-2 strain SB5 Cont	8.13e+02
813	4	25.0	772 10	R13478	Human TLE-3.	8.13e+02
814	4	25.0	772 10	R51111	Human TLE-3.	8.13e+02
815	4	25.0	774 26	W53390	Flavobacterium odorat	8.13e+02
816	4	25.0	778 17	R94900	Human tastin.	8.13e+02
817	4	25.0	778 3	R13456	Duffy receptor.	8.13e+02
818	4	25.0	783 24	W27113	Rat spleen pro-hormon	8.13e+02
819	4	25.0	784 27	W34179	Human GC binding prot	8.13e+02



820	4	25.0	790 34	W59808	Partial amino acid se	8.13e+02	893	4	25.0	1042 14	R70122	IL18-R type 1-GBP 130	8.13e+02
821	4	25.0	797 38	W85042	PKB-green fluorescent	8.13e+02	894	4	25.0	1045 38	W85050	Stat5-green fluoresce	8.13e+02
822	4	25.0	797 38	W85043	NFKappaB p65 subunit-	8.13e+02	895	4	25.0	1047 5	R11137	GAP6 encoded by lambd	8.13e+02
823	4	25.0	798 38	W85025	Smad4-green fluoresce	8.13e+02	896	4	25.0	1047 5	R25336	Lambda clone 101 prot	8.13e+02
824	4	25.0	805 36	W85024	HSV-2 strain S85 Cont	8.13e+02	897	4	25.0	1047 2	R06328	Sequence of full leng	8.13e+02
825	4	25.0	806 38	W85013	Smad4-green fluoresce	8.13e+02	898	4	25.0	1052 28	W37412	Human G-protein coupl	8.13e+02
826	4	25.0	814 2	R12157	Recombinant collagena	8.13e+02	899	4	25.0	1061 13	R70231	P. vivax DABP	8.13e+02
827	4	25.0	816 13	R66931	AMML chromosome inv(1	8.13e+02	900	4	25.0	1064 14	R70231	IL18-R type 2-GBP 130	8.13e+02
828	4	25.0	816 13	R39647	Human fibroblast grow	8.13e+02	901	4	25.0	1065 30	W33819	Arabidopsis cellulose	8.13e+02
829	4	25.0	818 36	R47233	HSV-2 strain S85 Cont	8.13e+02	902	4	25.0	1070 23	W17789	Green fluorescent pro	8.13e+02
830	4	25.0	820 9	R47233	Human fibroblast grow	8.13e+02	903	4	25.0	1071 34	W49095	Human inhibitory fact	8.13e+02
831	4	25.0	821 36	W72250	HSV-2 strain S85 Cont	8.13e+02	904	4	25.0	1107 18	W01312	Human DNA polymerase	8.13e+02
832	4	25.0	821 7	R35451	Mouse eps8.	8.13e+02	905	4	25.0	1115 3	R13457	Duffy receptor.	8.13e+02
833	4	25.0	822 5	R26337	N-sam.	8.13e+02	906	4	25.0	1115 23	R22478	Duffy antigen binding	8.13e+02
834	4	25.0	822 32	W56490	Zea mays starch branc	8.13e+02	907	4	25.0	1122 21	R14052	TIM protein.	8.13e+02
835	4	25.0	822 3	R13549	Basic FGF receptor.	8.13e+02	908	4	25.0	1140 16	R81731	GST-SEP fusion protei	8.13e+02
836	4	25.0	823 17	R93800	Arther earl gene prod	8.13e+02	909	4	25.0	1147 8	R41199	CAI antigen.	8.13e+02
837	4	25.0	836 38	W85017	Grk5-green fluorescen	8.13e+02	910	4	25.0	1171 38	W85037	Green fluorescent pro	8.13e+02
838	4	25.0	842 38	W85008	Grk5-green fluorescen	8.13e+02	911	4	25.0	1178 21	W20174	H. pylori cytoplasmic	8.13e+02
839	4	25.0	843 38	W85030	Green fluorescent pro	8.13e+02	912	4	25.0	1181 38	W85038	NFAT1-green fluoresce	8.13e+02
840	4	25.0	844 23	W19213	Corn starch branching	8.13e+02	913	4	25.0	1181 13	R72593	H. pylori tagA antige	8.13e+02
841	4	25.0	847 39	W85603	Hexosaminidase enzyme	8.13e+02	914	4	25.0	1183 22	W20884	H. pylori cytoplasmic	8.13e+02
842	4	25.0	848 34	W25677	Homo sapiens clone CS	8.13e+02	915	4	25.0	1198 33	W64384	S. cerevisiae L9470.2	8.13e+02
843	4	25.0	850 18	R91308	Helicobacter pylori T	8.13e+02	916	4	25.0	1224 28	W26735	Staphylococcus carnos	8.13e+02
844	4	25.0	853 38	W85031	Protein tyrosine phos	8.13e+02	917	4	25.0	1245 13	R70106	TNF-R-PI. vivax Duffy	8.13e+02
845	4	25.0	859 3	R12893	OmpAL streptokinase-s	8.13e+02	918	4	25.0	1307 19	R99255	Aspergillus flavus mu	8.13e+02
846	4	25.0	859 13	R72594	H. pylori tagA antige	8.13e+02	919	4	25.0	1312 36	W22775	Human RAD50.	8.13e+02
847	4	25.0	862 28	R31655	Pacific yew taxadiene	8.13e+02	920	4	25.0	1312 35	W71295	Human homologue of ye	8.13e+02
848	4	25.0	865 38	W85027	zap70-green fluoresce	8.13e+02	921	4	25.0	1338 37	W80600	Helicobacter pylori C	8.13e+02
849	4	25.0	868 2	R51143	Porcine Factor VIII.	8.13e+02	922	4	25.0	1370 3	P60005	Sequence encoded by h	8.13e+02
850	4	25.0	868 8	R07454	Second open reading f	8.13e+02	923	4	25.0	1382 15	R77440	Wild type human insul	8.13e+02
851	4	25.0	870 11	R59324	Human GAP protein.	8.13e+02	924	4	25.0	1389 21	W14051	TIM protein splice va	8.13e+02
852	4	25.0	871 38	W85026	Green fluorescent pro	8.13e+02	925	4	25.0	1405 31	W56101	Enhanced green fluore	8.13e+02
853	4	25.0	872 20	R96039	95 kD protein.	8.13e+02	926	4	25.0	1416 13	R67358	Human astrovirus sero	8.13e+02
854	4	25.0	874 1	P82981	SP6 bacteriophage RNA	8.13e+02	927	4	25.0	1435 23	W22477	Silaic acid binding p	8.13e+02
855	4	25.0	881 28	W47124	S. cerevisiae transcr	8.13e+02	928	4	25.0	1435 13	R70232	P. falciparum SAMP.	8.13e+02
856	4	25.0	882 10	R51930	Mutated GAL4 protein	8.13e+02	929	4	25.0	1452 9	R42471	Canine coronavirus 1-	8.13e+02
857	4	25.0	885 13	R66930	AMML chromosome inv(1	8.13e+02	930	4	25.0	1453 6	R31038	E2 protein of canine	8.13e+02
858	4	25.0	886 39	W89787	Staphylococcus aureus	8.13e+02	931	4	25.0	1454 5	R24397	Prod. of the S gene o	8.13e+02
859	4	25.0	889 10	R56248	Xenopus thrombospondi	8.13e+02	932	4	25.0	1454 9	R42474	FEVCF/FIPV chimeric sp	8.13e+02
860	4	25.0	892 19	W05139	scFv2(FRP5/225)-ETA (	8.13e+02	933	4	25.0	1454 9	R42464	Feline infectious per	8.13e+02
861	4	25.0	892 19	W05140	scFv2(FRP5/225)-ETA (	8.13e+02	934	4	25.0	1454 9	R42478	FEVCF/FIPV chimeric sp	8.13e+02
862	4	25.0	898 3	P61082	Complete translation	8.13e+02	935	4	25.0	1454 9	R42472	FEVCF/FIPV chimeric sp	8.13e+02
863	4	25.0	898 3	P61030	Entire coded sequence	8.13e+02	936	4	25.0	1454 5	R24511	Prod. of the S genome	8.13e+02
864	4	25.0	899 3	P61056	Translation of plasmid	8.13e+02	937	4	25.0	1454 9	R42467	Feline infectious per	8.13e+02
865	4	25.0	900 27	W33887	Flea saliva protein p	8.13e+02	938	4	25.0	1454 9	R42477	FEVCF/FIPV chimeric sp	8.13e+02
866	4	25.0	908 29	W41568	Human metabotropic gl	8.13e+02	939	4	25.0	1454 9	R42470	Feline infectious per	8.13e+02
867	4	25.0	908 29	W49288	Human metabotropic gl	8.13e+02	940	4	25.0	1454 9	R42475	FEVCF/FIPV chimeric sp	8.13e+02
868	4	25.0	915 31	W52306	Human pro-protein con	8.13e+02	941	4	25.0	1454 5	R24395	Prod. of the S gene o	8.13e+02
869	4	25.0	916 38	W85023	Amino acid sequence o	8.13e+02	942	4	25.0	1470 28	W23411	Porcine transmissible	8.13e+02
870	4	25.0	932 38	W85040	PKG-green fluorescent	8.13e+02	943	4	25.0	1477 8	R41728	High molecular weight	8.13e+02
871	4	25.0	932 30	W53281	Murine phospholipase	8.13e+02	944	4	25.0	1477 27	W30294	Non-typeable Haemophi	8.13e+02
872	4	25.0	933 38	W85039	Green fluorescent pro	8.13e+02	945	4	25.0	1477 8	R41724	High molecular weight	8.13e+02
873	4	25.0	936 39	W89801	Staphylococcus aureus	8.13e+02	946	4	25.0	1529 8	R41732	High molecular weight	8.13e+02
874	4	25.0	943 38	W83168	Human receptor tyrosi	8.13e+02	947	4	25.0	1594 5	P81183	Sequence of the peplo	8.13e+02
875	4	25.0	954 23	W19752	Yeast inhibitor of ap	8.13e+02	948	4	25.0	1601 27	R30292	Non-typeable Haemophi	8.13e+02
876	4	25.0	956 3	R15047	Soluble human insulin	8.13e+02	949	4	25.0	1604 13	R70105	TNF-R-EBA 175 fusion	8.13e+02
877	4	25.0	957 22	W20841	H. pylori transmembra	8.13e+02	950	4	25.0	1612 24	W24094	Ras-binding protein p	8.13e+02
878	4	25.0	961 10	R56249	Human thrombospondin-	8.13e+02	951	4	25.0	1612 12	R66457	Chimeric ALL-1/AF-6 p	8.13e+02
879	4	25.0	966 14	R76065	Yeast MSF2 protein.	8.13e+02	952	4	25.0	1614 38	W83312	Mouse Lrp5 protein.	8.13e+02
880	4	25.0	968 38	W85011	p85alpha-green floure	8.13e+02	953	4	25.0	1618 5	R27205	Human nestin.	8.13e+02
881	4	25.0	969 8	R41662	Paired basic amino ac	8.13e+02	954	4	25.0	1618 11	R60127	Human nestin protein	8.13e+02
882	4	25.0	970 38	W85020	p85alpha-green floure	8.13e+02	955	4	25.0	1664 1	P80264	Sequence of Heliothis	8.13e+02
883	4	25.0	971 33	W48896	Candida albicans CaCL	8.13e+02	956	4	25.0	1704 9	R49657	Neuroblastoma indicat	8.13e+02
884	4	25.0	982 22	W20980	H. Pylori secreted or	8.13e+02	957	4	25.0	1717 27	W23331	Cryptosporidium parvu	8.13e+02
885	4	25.0	985 21	W15191	Aspergillus oryzae al	8.13e+02	958	4	25.0	1721 31	W48299	Mouse DEC-205.	8.13e+02
886	4	25.0	990 7	R39343	EpIB protein.	8.13e+02	959	4	25.0	1723 19	W00645	Truncated plasmodium	8.13e+02
887	4	25.0	996 38	W85033	IkappaB kinase alpha	8.13e+02	960	4	25.0	1726 20	W00385	CD4-EBA175 fusion pro	8.13e+02
888	4	25.0	997 38	W85032	Green fluorescent pro	8.13e+02	961	4	25.0	1786 8	R41043	Microtubule-associate	8.13e+02
889	4	25.0	1020 19	W05141	scFv2(FRP5/225)-ETA (	8.13e+02	962	4	25.0	1831 27	W23329	Human calcium channel	8.13e+02
890	4	25.0	1028 8	R41044	Plasmodium vivax buff	8.13e+02	963	4	25.0	1854 35	W79161	Human calcium channel	8.13e+02
891	4	25.0	1039 38	W85014	Stats-green fluoresce	8.13e+02	964	4	25.0	1946 28	W47273	Lactobacillus bulgari	8.13e+02
892	4	25.0	1041 20	W11866	ATH1 gene product, At	8.13e+02	965	4	25.0	1972 20	W00024	Smooth muscle myosin	8.13e+02

966 4 25.0 2008 23 W22016 Utophin truncated po 8.13e+02  
967 4 25.0 2042 32 W56319 Haemophilus paraaelli 8.13e+02  
968 4 25.0 2133 29 W41133 Sus scrofa factor VII 8.13e+02  
969 4 25.0 2186 27 W31948 Mouse 22B/30B (candid 8.13e+02  
970 4 25.0 2273 20 R98811 Erysiphe graminis ace 8.13e+02  
971 4 25.0 2353 19 R93393 Haemophilus adhesio 8.13e+02  
972 4 25.0 2547 37 W81511 Drosophila fat facets 8.13e+02  
973 4 25.0 2549 32 W56027 FRAP (tor1) protein. 8.13e+02  
974 4 25.0 2549 19 W04235 Human RAP11 8.13e+02  
975 4 25.0 2555 37 W81512 Drosophila fat facets 8.13e+02  
976 4 25.0 2695 20 W07632 Human type I inositol 8.13e+02  
977 4 25.0 2713 19 W00168 Human T cell inositol 8.13e+02  
978 4 25.0 2749 3 R13887 Inositol-3-phosphate 8.13e+02  
979 4 25.0 2864 16 R82072 Hepatitis GB virus (H 8.13e+02  
980 4 25.0 2893 35 W71556 Helicobacter polyep 8.13e+02  
981 4 25.0 2913 20 W00384 Plasmodium falcipar 8.13e+02  
982 4 25.0 2938 11 R59223 GAP protein Iral. 8.13e+02  
983 4 25.0 2985 32 W56643 DNA-dependent protein 8.13e+02  
984 4 25.0 3077 1 P93283 Sequence of clone HIV 8.13e+02  
985 4 25.0 3144 26 W36887 Previously undescribe 8.13e+02  
986 4 25.0 3144 11 R58777 Protein encoded by Hu 8.13e+02  
987 4 25.0 3144 22 W09871 Human huntingtin. 8.13e+02  
988 4 25.0 3210 1 P81770 Deduced sequence enco 8.13e+02  
989 4 25.0 3457 12 R62504 Large polyprotein seq 8.13e+02  
990 4 25.0 3672 27 W31950 Human bg protein asso 8.13e+02  
991 4 25.0 3788 25 W21594 Murine Lys1 long iso 8.13e+02  
992 4 25.0 3801 27 W31949 Human bg protein asso 8.13e+02  
993 4 25.0 3960 32 W56642 Protein kinase cataly 8.13e+02  
994 4 25.0 4544 11 R60517 Human alpha-2-MR. 8.13e+02  
995 4 25.0 4544 9 R47861 Alpha 2-Macroglobulin 8.13e+02  
996 4 25.0 4866 37 W77410 Human ryanodin recept 8.13e+02  
997 4 25.0 4987 3 R10834 Rianodin receptor. 8.13e+02  
998 4 25.0 5035 5 R25450 MH mutant porcine rya 8.13e+02  
999 4 25.0 5072 2 R11510 Ryanodine receptor de 8.13e+02  
1000 4 25.0 15281 9 R44929 T. niveum Cyclosporin 8.13e+02

## ALIGNMENTS

RESULT 1  
ID W06548 standard; Protein; 228 AA.  
AC W06548;  
DT 13-MAR-1997 (first entry)  
DE Human colon specific gene CSG5 polypeptide fragment.  
KW Colon specific gene; CSG5; colon cancer; metastasis; diagnosis;  
KW therapy; antibody; vaccine.  
OS Homo sapiens.  
PN W09639419-A1.  
PD 12-DEC-1996.  
PF 06-JUN-1995; U07289.  
PR 06-JUN-1995; WO-U07289.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Yu G;  
DR WPI; 97-043054/04.  
DR N-PSDB; T45884.  
PT Human colon specific genes and their expression products - detection  
PT of which, in non-colon tissue samples, can be used as indication of  
PT colon cancer metastasis  
PS Claim 8; Fig 5; 60pp; English.  
CC Novel polypeptides (W06545-53) are encoded by cDNA clones (see also  
CC T45880-92) corresponding to 13 human colon specific genes, in  
CC designated CSG1, CSG2, etc., that are primarily expressed in  
CC tissues derived from the colon. Recombinant CSG polypeptides can  
CC be produced in transformed host cells. They are useful diagnostic  
CC markers for colon cancer and for colon cancer metastasis and can  
CC also be used to screen for (ant)agonist cpds. of therapeutic or  
CC diagnostic value. Antibodies raised against the colon-specific  
CC polypeptides may be used to target colon cancer cells or as part  
CC of a colon cancer vaccine.  
SQ Sequence 228 AA;

Query Match 100.0%; Score 16; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.54e-13;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 59 einkddvqhkvqcfsr 74  
Qy 1 EINKDDVQHVKQVCFSR 16  
RESULT 2  
ID W46879 standard; Protein; 228 AA.  
AC W46879;  
DT 22-JUN-1998 (first entry)  
DE Protein sequence encoded by a colon-specific gene.  
KW Colon-specific gene; probe; detection; expression; human;  
KW diagnostic assay; colon cancer; antibody; screening.  
OS Homo sapiens.  
PN US5733748-A.  
PD 31-MAR-1998.  
PF 06-JUN-1995; 469667.  
PR 06-JUN-1995; US-469667.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen C, Yu G;  
DR WPI; 98-229823/20.  
DR N-PSDB; V16672.  
PT Colon-specific nucleic acids - useful as probes for detecting colon  
PT cancer micrometastases  
PS Claim 1; Fig 5; 51pp; English.  
CC W46876-80 and W4682-85 represent proteins encoded by colon-specific  
CC genes. The polynucleotides encoding these proteins can be used  
CC as probes to detect expression of the corresponding human genes,  
CC e.g. in diagnostic assays for detecting micrometastases of colon cancer.  
CC Recombinant cells containing the polynucleotides can be used to  
CC produce the proteins, in order that antibodies can be raised and  
CC used in further screening or diagnostics.  
SQ Sequence 228 AA;  
Query Match 100.0%; Score 16; DB 29; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2.54e-13;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 59 einkddvqhkvqcfsr 74  
Qy 1 EINKDDVQHVKQVCFSR 16  
RESULT 3  
ID R82657 standard; Protein; 877 AA.  
AC R82657;  
DT 20-DEC-1995 (first entry)  
DE Human mGluR3.  
KW Metabotropic glutamate receptor 3; mGluR3; stroke; epilepsy;  
KW Alzheimer disease; detection; diagnosis; therapy.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT domain 575..597  
FT /label= TMD-I  
FT /note= "transmembrane domain I"  
FT domain 612..632  
FT /label= TMD-II  
FT /note= "transmembrane domain II"  
FT domain 644..662  
FT /label= TMD-III  
FT /note= "transmembrane domain III"  
FT domain 686..707  
FT /label= TMD-IV  
FT /note= "transmembrane domain IV"  
FT domain 733..754  
FT /label= TMD-V  
FT /note= "transmembrane domain V"  
FT domain 768..789  
FT /label= TMD-VI  
FT /note= "transmembrane domain VI"  
FT domain 802..826  
FT /label= TMD-VII

/note= "transmembrane domain VII"

FT W09522609-A2.  
 PN 24-AUG-1995.  
 PF 21-FEB-1995; G00356.  
 PR 21-FEB-1994; GB-003285.  
 PR 01-AUG-1994; GB-015532.  
 PA (WELL.) WELLCOME FOUND LTD.  
 PI Makoff AJ;  
 DR WPI: 95-302715/39.  
 DR N-PSDB; T03887.  
 PT New isolated human metabotropic glutamate receptors - used for  
 PT detection, diagnosis and therapy of diseases associated with the  
 PT receptors, eg. stroke, epilepsy and Alzheimer's disease.  
 PS Claim 2; page 32-35; 55pp; English.  
 CC mRNA from the human brain median frontal cortex and amygdala was used  
 CC to construct a cDNA library in lambda ZAP II. cDNA was amplified by  
 CC PCR primers (T03889-95) based on rat mGluR1-5 sequences to obtain  
 CC cDNA encoding human mGluR3.  
 SQ Sequence 877 AA;

Query Match 37.5%; Score 6; DB 14; Length 877;  
 Best Local Similarity 100.0%; Pred. No. 4.57e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 76 einkdd 81  
 |||||  
 Qy 1 EINKDD 6

RESULT 4  
 ID R64252 standard; Protein: 879 AA.  
 AC R64252;  
 DT 21-JUL-1995 (first entry)  
 DE Human mGluR3.  
 KW Metabotropic glutamate receptor; mGluR3; hippocampus; cerebellum;  
 KW pCMV-T7-3; pCMV-T7-2.  
 OS Homo sapiens.  
 PN W09429449-A.  
 PD 22-DEC-1994.  
 PF 03-JUN-1994; U06273.  
 PR 04-JUN-1993; US-072574.  
 PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.  
 PI Daggett L, Ellis SB, Hess SD, Johnson EC, Liaw C;  
 PI Pontsler A;  
 DR WPI: 95-036478/05.  
 DR N-PSDB; Q80418.  
 PT New DNA encoding human metabotropic glutamate receptor  
 PT sub-type(s) - and related proteins, probes, RNA, transformed  
 PT cells and antibodies, useful in treatment, diagnosis and  
 PT identification of specific modulators  
 PS Claim 11; Page 83-87; 125pp; English.  
 CC Human hippocampus and cerebellum cDNA libraries were screened with  
 CC DNA encoding the rat mGluR2 receptor. Several clones were isolated  
 CC and these were ligated to obtain a full-length construct (given in  
 CC Q80418) encoding human mGluR3 (R64252). The mGluR3 coding sequence  
 CC is operatively linked to the regulatory elements of pCMV-T7-3 or  
 CC pCMV-T7-2 for expression in mammalian cells.  
 SQ Sequence 879 AA;

Query Match 37.5%; Score 6; DB 12; Length 879;  
 Best Local Similarity 100.0%; Pred. No. 4.57e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 einkdd 83  
 |||||  
 Qy 1 EINKDD 6

RESULT 5  
 ID W82486 standard; Protein: 159 AA.  
 AC W82486;  
 DT 02-MAR-1999 (first entry)  
 DE Ehrlichia sp. E74.1 protein.

KW Granulocytic ehrlichia; GE; E74; tick-borne infection; fatal; vaccine;  
 KW immune response; detection; diagnosis; Ehrlichiosis.  
 OS Ehrlichia sp.  
 PN W09849312-A2.  
 PD 05-NOV-1998.  
 PF 24-APR-1998; U08264.  
 PR 25-APR-1997; US-044869.  
 PA (AQUI-) AQUILA BIOPHARMACEUTICALS INC.  
 PI Beitz G, Coughlin RT, Murphy C, Storey J;  
 DR WPI: 99-034663/03.  
 DR N-PSDB; V65141.  
 PT New isolated granulocytic ehrlichia nucleic acids - used to develop  
 PT products for use in vaccines for inhibiting Ehrlichiosis and for use  
 PT in detection and diagnosis  
 PS Disclosure; Fig 8; 184pp; English.  
 CC This sequence represents a protein, E74.1 encoded by ORF1 of an  
 CC Ehrlichia sp. which is associated with granulocytic ehrlichia (GE),  
 CC and has been isolated from HL60 cells infected with Ehrlichia sp. GE  
 CC is an acute potentially fatal tick borne infection and the proteins  
 CC described in this invention can be used in vaccines to elicit a  
 CC beneficial immune response in an animal to GE. They can be used for  
 CC inhibiting Ehrlichiosis in an animal. The products can also be used for  
 CC detection and diagnosis.  
 SQ Sequence 159 AA;

Query Match 31.3%; Score 5; DB 38; Length 159;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 149 einkd 153  
 |||||  
 Qy 1 EINKD 5

RESULT 6  
 ID R46497 standard; Protein: 238 AA.  
 AC R46497;  
 DT 09-JUL-1994 (first entry)  
 DE Human GAP-43.  
 KW GAP-43; internal regulatory protein; IRP; recombinant; neuron;  
 KW neuronal growth.  
 OS Homo sapiens.  
 PN A09347435-A.  
 PD 23-DEC-1993.  
 PF 17-SEP-1993; 047435.  
 PR 22-DEC-1989; US-465635.  
 PR 02-JUL-1990; US-546453.  
 PA (FEDE/) FEDEROFF H J.  
 PA (FISH/) FISHMAN M C.  
 PA (STRI/) STRITTMATTER SM.  
 PA (VALE/) VALENZUELA D.  
 PA (ZUBE/) ZUBER M X.  
 PI Federoff HJ, Fishman MC, Strittmatter SM, Valenzuela D;  
 PI Zuber MX;  
 DR WPI: 94-049278/07.  
 DR P-PSDB; Q57503.  
 PT Recombinant mammalian GAP-43 protein - used to monitor and  
 PT regulate neuronal growth in animals, pref. humans  
 PS Disclosure; Fig 5A; 156pp; English.  
 CC The nucleotide and corresponding amino acid sequences for human GAP-43  
 CC (Q57503, R46497) and rat GAP-43 (Q57502, R46496) were determined.  
 CC Such sequences may be expressed in prokaryotic or eukaryotic hosts.  
 SQ Sequence 238 AA;

Query Match 31.3%; Score 5; DB 9; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 kddvq 60  
 |||||  
 Qy 4 KDDVQ 8

RESULT 7  
 ID R07038 standard; protein; 238 AA.  
 AC R07038;  
 DT 26-NOV-1990 (first entry)  
 DE Human neuronal growth related protein, GAP-43.  
 KW Internal regulatory protein; IRP; neuronal growth related peptide;  
 KW GAP-43; membrane targeting peptide; MTP; ischaemia; hypoxia; stroke.  
 OS Homo sapiens.  
 PN W09006948-A.  
 PD 28-JUN-1990.  
 PF 22-DEC-1989; U05808.  
 PR 22-DEC-1988; US-288604.  
 PR 02-FEB-1989; US-305239.  
 PR 01-SEP-1989; US-401408.  
 PA (FISH/) FISHMAN M C.  
 PI FISHMAN MC, ZUBER MX, STRITTMAT SM, VALENZUELA D;  
 DR WPI; 90-224493/29.  
 DR N-PSDB; Q05302.  
 PT Mammalian neuronal growth peptide GAP-43 and corresp. DNA - also  
 PT new membrane targeting and internal regulatory peptide(s),  
 PT useful e.g. for neuronal modelling and healing neural tissue  
 PT damage.  
 PS Disclosure; 9pp; English.  
 CC GAP-43 is useful in structural remodelling of neural cells and  
 CC treatment of ischaemia, hypoxia, stroke and other neural tissue  
 CC damage.  
 SQ Sequence 238 AA;  
 Query Match 31.3%; Score 5; DB 1; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 56 kddvq 60  
 I I I I I  
 QY 4 KDDVQ 8

RESULT 8  
 ID R25333 standard; Protein; 238 AA.  
 AC R25333;  
 DT 03-APR-1993 (first entry)  
 DE Sequence encoded by human GAP-43 CDNA.  
 KW Neuronal growth-related protein; axon; GAP-43 protein.  
 OS Homo sapiens.  
 PN W09218138-A.  
 PD 29-OCT-1992.  
 PF 10-APR-1992; U03014.  
 PR 10-APR-1991; US-683455.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PI Federoff HJ, Fishman MC, Strittmatter SM, Valenzuela D,  
 PI Zuber MX;  
 DR WPI; 92-381769/46.  
 DR N-PSDB; Q30681.  
 PT Recombinant mammalian GAP-43 - used in the study, diagnosis and  
 PT treatment of damaged, diseased or dysfunctional nervous system  
 PT tissue  
 PS Claim 5; Fig 5A; 191pp; English.  
 CC A cDNA library was generated from RNA of rat dorsal root ganglion  
 CC from embryonic day 17 rats and cloned into the lambda gtl1  
 CC expression vector. Three presumptive GAP-43 clones were identified  
 CC with antibody to GAP-43. The complete nucleotide sequence for the  
 CC longest clone, GAP-43-2 was detd. The predicted mol. size of the  
 CC GAP-43 protein from the open reading frame is 24 kD. The rat GAP-43  
 CC cDNA was labelled and used as a probe to isolate human GAP-43 CDNA  
 CC from a human brain cDNA library. The longest clone is Cla from the  
 CC cerebellum library (Q30681). GAP-43 is identical to a neural  
 CC specific mouse calmodulin-binding protein, termed p-57. The proteins  
 CC are highly conserved between human, rat and mouse.  
 SQ Sequence 238 AA;  
 Query Match 31.3%; Score 5; DB 6; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 56 kddvq 60  
 I I I I I  
 QY 4 KDDVQ 8

RESULT 9  
 ID R30016 standard; Protein; 238 AA.  
 AC R30016;  
 DT 27-APR-1993 (first entry)  
 DE Human GAP-43.  
 KW Acute ischemic injury; investigation; medical composition; compsn.  
 OS Homo sapiens.  
 PN J04066595-A.  
 PD 02-MAR-1992.  
 PF 26-OCT-1990; 290615.  
 PR 02-JUL-1990; US-546453.  
 PA (GEHO-) GEN HOSPITAL CORP.  
 DR WPI; 92-409475/50.  
 DR N-PSDB; Q32849.  
 PT Mammal GAP-43 compsn. used to regulate and control acute ischemic  
 PT injuries - is prepd. from host cell transfected with vector of  
 PT cDNA that codes mammalian GAP-43  
 PS Claim 1; Fig 5; 13pp; Japanese.  
 CC The sequence is that of human GAP-43 which may be recombinantly  
 CC produced and used in medical compsn. with carriers. It may  
 CC also be used in a compsn. used to regulate and control acute  
 CC ischemic injuries. Anti GAP-43 monoclonal antibodies can be raised  
 CC and when labelled, used to detect/measure GAP-43 in a sample.  
 SQ Sequence 238 AA;  
 Query Match 31.3%; Score 5; DB 6; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 56 kddvq 60  
 I I I I I  
 QY 4 KDDVQ 8

RESULT 10  
 ID R46486 standard; Protein; 238 AA.  
 AC R46486;  
 DT 09-JUL-1994 (first entry)  
 DE Human GAP-43.  
 KW GAP-43; internal regulatory protein; IRP; recombinant; neuron;  
 KW neuronal growth.  
 OS Homo sapiens.  
 PN AU9062068-A.  
 PD 03-JAN-1991.  
 PF 31-AUG-1990; 062068.  
 PR 22-DEC-1989; US-465635.  
 PR 02-JUL-1990; US-546453.  
 PA (FEDE/) FEDEROFF H J.  
 PA (FISH/) FISHMAN M C.  
 PA (STRI/) STRITTMATTER SM.  
 PA (VALE/) VALENZUELA D.  
 PA (ZUBE/) ZUBER M X.  
 PI Federoff HJ, Fishman MC, Strittmatter SM, Valenzuela D;  
 PI Zuber MX;  
 DR WPI; 94-049230/07.  
 DR P-PSDB; Q57496.  
 PT Recombinant mammalian GAP-43 protein - used to monitor and  
 PT regulate neuronal growth in animals, pref. humans  
 PS Disclosure; Fig 5A; 156pp; English.  
 CC The nucleotide and corresponding amino acid sequences for human GAP-43  
 CC (Q57496, R46486) and rat GAP-43 (Q57495, R46485) were determined.  
 CC Such sequences may be expressed in prokaryotic or eukaryotic hosts.  
 SQ Sequence 238 AA;  
 Query Match 31.3%; Score 5; DB 9; Length 238;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 56 kdqvq 60
   |||||
Qy 4 KDDVQ 8

RESULT 11
ID W64549 standard; Protein: 327 AA.
AC W64549;
DE Human fibrosarcoma cell line HT-1080 clone HP10196 protein.
KW Transmembrane domain; human; nutrition; cytokine; cell proliferation;
KW differentiation; immune system; stimulator; suppressor; regulator;
KW hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;
KW haemostatic; thrombolytic; ligand; anti-inflammatory; tumour.
OS Homo sapiens.
PN W09821328-A2.
PD 22-MAY-1998.
PF 07-NOV-1997; J04056.
PR 13-NOV-1996; JP-301429.
PA (PROT-) PROTEGENE INC.
PI (SAGA) SAGAMI CHEM RES CENTRE.
PI Kato S, Kobayashi M, Sekine S, Yamaguchi T;
DR WPI: 98-297932/26.
DR N-PSDB: V49580, V49581.
PT Human protein having transmembrane domain - useful for, e.g.
PT research and nutrition
PS Claim 1: Page 107-108; 205pp; English.
CC W64534-W64558 represent human proteins containing a transmembrane domain.
CC These proteins can be used for, e.g. research and nutrition, and may have
CC cytokine and cell proliferation/differentiation, immune
CC stimulating/suppressing, haematopoiesis regulating, tissue growth,
CC activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,
CC receptor/ligand, anti-inflammatory or tumour inhibition activity.
SQ Sequence 327 AA;

Query Match 31.3%; Score 5; DB 34; Length 327;
Best Local Similarity 100.0%; Pred. No. 6.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 139 hkqvc 143
   |||||
Qy 9 HKQVC 13

RESULT 12
ID W76830 standard; Protein: 330 AA.
AC W76830;
DE Human GRBP protein.
KW Growth factor receptor binding protein; GRBP; human; therapy; treatment;
KW sympathetic nervous system; hypertension; myocardial infarction; angina;
KW cancer; inflammation; immune response; screening; diagnosis; modulator;
KW gene mapping.
OS Homo sapiens.
PN W09840482-A1.
PD 17-SEP-1998.
PF 09-MAR-1998; U04529.
PR 11-MAR-1997; US-815176.
PA (INCY-) INCYTE PHARM INC.
PI Bandman O, Diegido AP;
DR WPI: 98-531523/45.
DR N-PSDB: V63097.
PT Human growth factor receptor binding protein, GRBP - useful e.g. to
PT treat sympathetic nervous system disorders and screen for
PT antagonists useful to treat cancer or immunological disorders
PS Claim 1; Fig 1A-E; 61pp; English.
CC This sequence represents a novel human growth factor receptor binding
CC protein (GRBP). GRBP (or its fragments) can be administered
CC therapeutically to treat disorders of the sympathetic nervous system,
CC e.g. hypertension, myocardial infarction, angina, etc. GRBP occurs
CC in cancerous tissues and tissues associated with the sympathetic nervous
CC system or with inflammation and the immune response. Increased expression

```

was proposed to be associated with cancer and immune disorders, while decreased expression/activity may be associated with sympathetic nervous system disorders. GRBP can be used to screen for antagonists and agonists, and the antagonists administered to treat cancer (e.g. adenocarcinoma, sarcoma, melanoma etc.) or immunological disorders (e.g. bronchitis, asthma, multiple sclerosis, etc.). It can be used to generate antibodies which can be used to diagnose conditions/diseases characterised by GRBP expression and to monitor therapeutic interventions. Polynucleotides encoding GRBP, or complementary sequences can be used as hybridisation probes or can produce complementary or antisense sequences for therapeutic administration to modulate/prevent GRBP expression, e.g. to treat/prevent cancer or immunological disorders. They can also be used to detect related sequences or for gene mapping.

Query Match 31.3%; Score 5; DB 37; Length 330;  
Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 100 ddvqh 104  
|||||  
Qy 5 DDVQH 9

RESULT 13

ID R24298 standard; Protein: 343 AA.

AC R24298;  
DE 20-NOV-1992 (first entry)  
KW D-alanine-D-alanine ligase VanA from E.faecium.  
KW Glycopeptide antibiotic; vancomycin; teicoplanin; resistant;  
KW D-Ala-D-Ala ligase; peptidoglycan precursor; transposon;  
KW inverted repeats.  
OS Enterococcus faecium BM4147.  
PN W09207942-A.  
PD 14-MAY-1992.  
PF 29-OCT-1991; F00855.  
PR 31-OCT-1990; FR-013579.  
PA (INSP) INST PASTEUR.  
PI Arthur M, Courvalin P, Dutka-malen S, Molinas C;  
DR WPI: 92-183677/22.  
DR N-PSDB: Q25178.  
PT Polypeptides involved in expression of glycopeptide antibiotic resistance - useful in diagnosing presence of Gram-positive enterococcal strains e.g. Enterococcus faecium and E.gallinarum Claim 1; Fig 8; 163pp; French.  
CC VanA is coded for by a 7.3kb HindIII-EcoRI fragment of the plasmid pIP816. (The plasmid contains resistance genes from Enterococcus faecium BM 4147 and was described in New England J. Med., 319:157-161). The VanA protein has a mol. mass of 37400.  
CC VanA interacts with VanH, a protein with similarity to D-specific alpha-keto acid reductases.. Resistance to glycopeptides is thought to involve the biosynthesis of a new component by VanA; the new peptide can be incorporated with peptidoglycans but is not recognised by vancomycin. See also Q25178-Q25183.  
SQ Sequence 343 AA;

Query Match 31.3%; Score 5; DB 2; Length 343;  
Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 inkdd 157  
|||||  
Qy 2 INKDD 6

RESULT 14

ID R11068 standard; Protein: 346 AA.

AC R11068;  
DE 23-MAY-1991 (first entry)  
KW 12D3 antigen sequence deduced from cDNA of pBb51.  
KW Babesia; red water fever; cattle; antibodies.  
OS Babesia bovis.  
FH Key Location/Qualifiers

FT peptide 1..21 /label= signal sequence  
 FT protein 22..346 /label= mature protein  
 FT  
 PN EP-417524-A.  
 PD 20-MAR-1991.  
 PF 23-AUG-1990; 116165.  
 PR 23-AUG-1989; AU-005902.  
 PA (CSIR ) COMMONWEALTH SCIENT ORG.  
 PI Riddles PW, Aylward JH, Wright IG;  
 DR WPI: 91-081822/12.  
 DR N-PSDB; Q10951.  
 PT New antigen from Babesia for protective vaccine - used with  
 PT derived antibodies, DNA sequences and oligo:nucleotide probes,  
 PT for immuno-diagnosis.  
 PS Claim 12; Fig 7; 40pp; English.  
 CC The sequence was deduced from a clone isolated from a cDNA library  
 CC prep'd. from bovine erythrocytes infected with Babesia bovis. The  
 CC DNA can be used to construct a vector for expression of the  
 CC antigen which can be used in vaccines to protect cattle from red  
 CC water fever.  
 CC See also R11069.  
 SQ Sequence 346 AA;

Query Match 31.3%; Score 5; DB 2; Length 346;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 242 einkd 246  
 Qy 1 EINKD 5

RESULT 15  
 ID W20686 standard; Protein; 350 AA.  
 AC W20686;  
 DT 14-JUL-1997 (first entry)  
 DE H. pylori cytoplasmic protein 04gell1713orf37.  
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope;  
 KW amino acid metabolism.  
 OS Helicobacter pylori.  
 PN W09640893-A1.  
 PD 19-DEC-1996.  
 PF 06-JUN-1996; U09122.  
 PR 07-JUN-1995; US-487032.  
 PR 01-APR-1996; US-630405.  
 PA (ASTR ) ASTRA AB.  
 PI Berglindh OT, Smith D, Mellgaerd BL;  
 DR WPI: 97-052306/05.  
 DR N-PSDB; T67939.  
 FT Helicobacter pylori nucleic acid sequences and related  
 FT polypeptide(s) - useful for vaccines to treat or prevent H. pylori  
 FT infection, and to detect Helicobacter  
 PS Claim 61; Pages 1106-1107; 1481pp; English.  
 CC The present sequence is a Helicobacter pylori cytoplasmic  
 CC protein involved in amino acid metabolism.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds,  
 CC useful as potential H. pylori life cycle activators or inhibitors.  
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from  
 CC overlapping contigs generated by mechanically shearing the bacterial  
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,  
 CC and the predicted coding regions defined by computer evaluation. To  
 CC identify likely H. pylori antigens for vaccine development, the amino  
 CC acid sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 SQ Sequence 350 AA;

Query Match 31.3%; Score 5; DB 22; Length 350;  
 Best Local Similarity 100.0%; Pred. No. 6.64e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 239 nkddv 243  
 Qy 3 NKDDV 7

Search completed: Sat Aug 28 15:14:45 1999  
 Job time : 73 secs.

\*\*\*\*\*  
 WAREHOUSE  
 \*\*\*\*\* (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993-1998 University of Edinburgh, U.K.  
 Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:12:26 1999; MasPar time 4.74 Seconds

Tabular output not generated. 135.354 Million cell updates/sec

Title: >US-09-049-696-49  
 Description: (1-16) from US09049696.pep  
 Perfect Score: 16  
 Sequence: 1 EINKDDVQHKQVCFSR 16

Scoring table: TABLE unitprotatable  
 Gap 50

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
 Listing first 1000 summaries

Database: pir60  
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 2.297; Variance 0.358; scale 6.417

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	37.5	183	2 B24479	calcium-binding prote	4.98e-01
2	6	37.5	183	2 A24479	calcium-binding prote	4.98e-01
3	6	37.5	185	2 S13183	calcium-binding prote	4.98e-01
4	6	37.5	185	2 S13184	calcium-binding prote	4.98e-01
5	6	37.5	185	2 S13182	calcium-binding prote	4.98e-01
6	6	37.5	185	2 S13185	calcium-binding prote	4.98e-01
7	6	37.5	440	2 B71153	hypothetical protein	4.98e-01
8	6	37.5	574	2 E64414	hypothetical protein	4.98e-01
9	6	37.5	671	2 A40692	signal recognition pa	4.98e-01
10	6	37.5	1881	2 S42369	Clathrin heavy chain	4.98e-01
11	6	37.5	1967	2 S21601	hypothetical protein,	4.98e-01
12	6	37.5	1968	1 PN0093	genome polyprotein -	4.98e-01
13	6	37.5	2703	2 A24420	notch protein - fruit	4.98e-01
14	5	31.3	60	2 A45681	hypothetical 7K prote	2.76e+01
15	5	31.3	107	2 E64356	conserved hypothetical	2.76e+01
16	5	31.3	133	2 B39002	probable membrane pro	2.76e+01
17	5	31.3	140	2 H64664	conserved hypothetical	2.76e+01
18	5	31.3	179	2 T00539	hypothetical protein	2.76e+01
19	5	31.3	183	2 G64348	precorrin-8W decarbox	2.76e+01
20	5	31.3	187	2 F65009	hypothetical protein	2.76e+01
21	5	31.3	219	2 G64490	triiose-phosphate isom	2.76e+01
22	5	31.3	228	2 S25730	sem-5 protein - Caeno	2.76e+01
23	5	31.3	238	2 I52638	neuronal growth-relat	2.76e+01

probable membrane pro 2.76e+01  
 conserved hypothetical 2.76e+01  
 hypothetical protein 2.76e+01  
 hypothetical protein 2.76e+01  
 probable lipoprotein 2.76e+01  
 beta-lactamase (EC 3. 2.76e+01  
 conserved hypothetical 2.76e+01  
 hypothetical protein 2.76e+01  
 APG5 protein - yeast 2.76e+01  
 rPIR protein - Escher 2.76e+01  
 probable proteinase s 2.76e+01  
 probable transcriptio 2.76e+01  
 hypothetical protein 2.76e+01  
 hypothetical protein 2.76e+01  
 hypothetical protein 2.76e+01  
 alkalal monooxygenase 2.76e+01  
 hypothetical protein 2.76e+01  
 DNA-directed RNA poly 2.76e+01  
 Grb-2 related adaptor 2.76e+01  
 hypothetical protein 2.76e+01  
 protochlorophyllide r 2.76e+01  
 protochlorophyllide r 2.76e+01  
 protochlorophyllide r 2.76e+01  
 lic-1 phase variation 2.76e+01  
 vancomycin resistance 2.76e+01  
 hypothetical protein 2.76e+01  
 AL1 protein - squash 2.76e+01  
 pectin lyase (EC 4.2. 2.76e+01  
 dopamine transporter 2.76e+01  
 probable membrane pro 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 NADH-dependent butano 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 hypothetical protein 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 phosphoprotein phosph 2.76e+01  
 nitrogen fixation pro 2.76e+01  
 hypothetical protein 2.76e+01  
 gene 24 protein - pha 2.76e+01  
 rad9 protein - Schiz 2.76e+01  
 N-acetylmuramoyl-L-al 2.76e+01  
 hypothetical protein 2.76e+01  
 aspartate transaminas 2.76e+01  
 conserved hypothetical 2.76e+01  
 vasoactive intestinal 2.76e+01  
 probable purine nucle 2.76e+01  
 55K erythrocyte membr 2.76e+01  
 probable probable n-a 2.76e+01  
 hypothetical protein 2.76e+01  
 farnesin - African claw 2.76e+01  
 Hms1 protein - yeast 2.76e+01  
 vasoactive intestinal 2.76e+01  
 ARS-binding protein - 2.76e+01  
 probable membrane pro 2.76e+01  
 dihydrolipoamide S-ac 2.76e+01  
 hypothetical protein 2.76e+01  
 secretion protein - A 2.76e+01  
 tryptophan 2-monooxyg 2.76e+01  
 Rab geranylgeranyl tr 2.76e+01  
 phosphoenolpyruvate-p 2.76e+01  
 probable membrane pro 2.76e+01  
 translation elongatio 2.76e+01  
 dopamine transporter 2.76e+01  
 dopamine transporter pr 2.76e+01  
 dopamine transporter 2.76e+01  
 dopamine transporter 2.76e+01  
 alpha 1,2 mannosidase 2.76e+01  
 penicillin-binding pr 2.76e+01  
 alpha 1,2 mannosidase 2.76e+01

239 2 S64327  
 243 2 E64452  
 267 2 E64172  
 276 2 E71705  
 279 2 S73526  
 281 1 PMSAP  
 284 2 H93355  
 289 2 S70293  
 32 2 S65160  
 33 2 S56317  
 34 2 E71697  
 308 2 S78291  
 309 2 T01554  
 314 2 T01554  
 322 2 E71137  
 325 2 G70334  
 326 2 S77954  
 327 2 T00797  
 328 2 B64095  
 330 3 J60376  
 331 2 E85076  
 333 2 S30915  
 333 2 B49850  
 333 2 S17823  
 333 2 A33465  
 339 1 C850VM  
 343 2 H64491  
 347 1 QCVS1  
 373 2 JH0155  
 385 2 I84455  
 387 2 S64082  
 389 2 S39781  
 390 2 S20392  
 390 2 S65672  
 390 2 S39780  
 390 2 B70011  
 393 2 I49016  
 396 2 A71206  
 397 2 J62524  
 406 2 A47492  
 408 2 S73681  
 418 2 A64508  
 419 2 S23425  
 427 1 GHP24  
 432 2 S41956  
 440 2 D64616  
 451 2 T00694  
 453 2 S56657  
 454 2 A64305  
 460 2 J62194  
 463 2 S37962  
 466 2 A39599  
 469 2 H71897  
 471 2 S05392  
 483 2 S40128  
 490 2 S28230  
 495 2 J62195  
 496 2 S33791  
 501 2 S54576  
 504 2 C42653  
 549 2 G71606  
 552 2 S55026  
 557 2 A25493  
 567 2 A45977  
 573 2 E70169  
 609 2 S67616  
 611 2 S38162  
 619 2 I59558  
 619 2 S20346  
 620 2 A48980  
 623 2 I57937  
 640 2 S60710  
 647 2 F70057  
 667 2 S60709

97	5	31.3	676	2	S65706	translation initiation	2.76e+01	170	4	25.0	112	2	E71167	hypothetical protein	9.32e+02
98	5	31.3	686	1	ALDYAT	amylase A (EC 3.2.1.-	2.76e+01	171	4	25.0	112	2	S00006	when protein - Arabia	9.32e+02
99	5	31.3	704	2	S39004	finger protein MSN2 -	2.76e+01	172	4	25.0	113	2	G69506	conserved hypothetical	9.32e+02
100	5	31.3	722	2	S64492	hypothetical protein	2.76e+01	173	4	25.0	113	2	S37883	hypothetical protein 2	9.32e+02
101	5	31.3	810	2	S59652	hypothetical protein	2.76e+01	174	4	25.0	114	2	I51230	gonadotropin alpha 2	9.32e+02
102	5	31.3	828	2	D71700	leucyl-tRNA synthetase	2.76e+01	175	4	25.0	117	2	I50992	gonadotropin alpha-su	9.32e+02
103	5	31.3	879	2	JH0562	metabolic glutamat	2.76e+01	176	4	25.0	117	2	S71325	DNA-directed RNA poly	9.32e+02
104	5	31.3	935	2	S63261	SEC21 protein - yeast	2.76e+01	177	4	25.0	117	2	A37198	glycoprotein hormones	9.32e+02
105	5	31.3	948	2	T03225	probable regulatory p	2.76e+01	178	4	25.0	118	2	A40554	glycoprotein hormones	9.32e+02
106	5	31.3	953	2	S66650	pyruvate,orthophospha	2.76e+01	179	4	25.0	119	2	A39426	saxiphilin - bullfrog	9.32e+02
107	5	31.3	953	2	S12894	pyruvate,orthophospha	2.76e+01	180	4	25.0	119	2	B25161	pillin protein traA pr	9.32e+02
108	5	31.3	959	2	S35061	protovillin - slime m	2.76e+01	181	4	25.0	119	2	I51229	gonadotropin alpha 1	9.32e+02
109	5	31.3	962	2	JH0669	helicase II-like prot	2.76e+01	182	4	25.0	120	2	UTPGA	gonadotropin alpha 1	9.32e+02
110	5	31.3	1118	2	S44641	F37A4 4 protein - Cae	2.76e+01	183	4	25.0	120	2	G00021	chlorionic gonadotropi	9.32e+02
111	5	31.3	1120	2	T01863	hypothetical protein	2.76e+01	184	4	25.0	122	2	A49773	ecdysone-dependent cu	9.32e+02
112	5	31.3	1208	2	T00362	hypothetical protein	2.76e+01	185	4	25.0	124	2	A53055	H+-transporting ATPase	9.32e+02
113	5	31.3	1222	2	S40977	hypothetical protein	2.76e+01	186	4	25.0	124	2	H69520	conserved hypothetical	9.32e+02
114	5	31.3	1224	2	S73171	DNA-directed RNA poly	2.76e+01	187	4	25.0	124	2	S67606	probable membrane pro	9.32e+02
115	5	31.3	1253	2	S46949	arylphorin receptor -	2.76e+01	188	4	25.0	125	2	JX0052	phospholipase A2 (EC	9.32e+02
116	5	31.3	1308	2	E71622	probable membrane ass	2.76e+01	189	4	25.0	127	2	C71152	hypothetical protein	9.32e+02
117	5	31.3	1311	2	A56390	mannosyl-glycoprotein	2.76e+01	190	4	25.0	130	2	D64559	hypothetical protein	9.32e+02
118	5	31.3	1317	2	S77517	DNA-directed RNA poly	2.76e+01	191	4	25.0	133	2	S33096	J5L protein - variola	9.32e+02
119	5	31.3	1358	1	XOCHDH	xanthine dehydrogenas	2.76e+01	192	4	25.0	133	2	H42513	J5L protein - vaccini	9.32e+02
120	5	31.3	1417	2	T00661	hypothetical protein	2.76e+01	193	4	25.0	134	2	S42040	hypothetical protein	9.32e+02
121	5	31.3	1447	2	S63669	UDPglucose--glycoprot	2.76e+01	194	4	25.0	134	2	C711348	hypothetical protein	9.32e+02
122	5	31.3	1481	2	A47341	amylopullulanase prec	2.76e+01	195	4	25.0	138	2	E71093	hypothetical protein	9.32e+02
123	5	31.3	1481	2	S28669	amylopullulanase prec	2.76e+01	196	4	25.0	139	2	G70686	hypothetical protein	9.32e+02
124	5	31.3	1714	2	E71609	Ser/Thr protein kinase	2.76e+01	197	4	25.0	140	2	H71197	hypothetical protein	9.32e+02
125	5	31.3	2013	2	C71610	probable membrane ass	2.76e+01	198	4	25.0	140	2	S14238	Ig gamma-1 chain prec	9.32e+02
126	5	31.3	2388	2	JH0271	beta spectrin, beta_5	2.76e+01	199	4	25.0	141	2	H64361	hypothetical protein	9.32e+02
127	5	31.3	2843	1	RHUWAP	adenomatous polyposis	2.76e+01	200	4	25.0	142	2	A46523	Charcot-Leyden crysta	9.32e+02
128	5	31.3	2845	2	I49505	adenomatous polyposis	2.76e+01	201	4	25.0	142	2	A47142	cystatin D precursor	9.32e+02
129	5	31.3	3169	2	T00296	toxin B - Escherichia	2.76e+01	202	4	25.0	144	2	S68454	SBB7 protein homolog	9.32e+02
130	5	31.3	4572	2	S57908	hypothetical polyprot	2.76e+01	203	4	25.0	146	1	QOV23	D2L protein - vaccini	9.32e+02
131	5	31.3	5037	2	A34161	ryanodine-binding pro	2.76e+01	204	4	25.0	146	2	S33106	D2L protein - variola	9.32e+02
132	4	25.0	19	2	S48488	succinyl-CoA syntheta	9.32e+02	205	4	25.0	147	2	H71016	hypothetical protein	9.32e+02
133	4	25.0	20	2	A60294	juvenile hormone-bind	9.32e+02	206	4	25.0	148	2	C71829	hypothetical protein	9.32e+02
134	4	25.0	32	2	G70249	hypothetical protein	9.32e+02	207	4	25.0	150	1	S49360	protein-tyrosine-phos	9.32e+02
135	4	25.0	38	2	A60583	glycoprotein hormones	9.32e+02	208	4	25.0	156	2	F69518	molybdenum cofactor b	9.32e+02
136	4	25.0	40	2	JH0756	sucrose-specific enzy	9.32e+02	209	4	25.0	157	2	S75510	hypothetical protein	9.32e+02
137	4	25.0	41	2	E84632	hypothetical protein	9.32e+02	210	4	25.0	157	2	S76488	hypothetical protein	9.32e+02
138	4	25.0	50	2	I53291	gene IGF-I protein -	9.32e+02	211	4	25.0	158	2	JC5297	vesicle-membrane fusi	9.32e+02
139	4	25.0	53	2	A61219	serpin Spi 1 - horse	9.32e+02	212	4	25.0	161	2	S69456	hypothetical protein	9.32e+02
140	4	25.0	56	2	S30243	Sry-type HMG box prot	9.32e+02	213	4	25.0	163	1	CFYCA	C-phycocyanin alpha c	9.32e+02
141	4	25.0	60	2	S14801	homeotic protein Hox	9.32e+02	214	4	25.0	163	2	F71635	hypothetical protein	9.32e+02
142	4	25.0	70	2	D56857	B2R protein - variol	9.32e+02	215	4	25.0	164	1	GCCHH	glycine cleavage syst	9.32e+02
143	4	25.0	72	1	TNLJH4	trans-activating tran	9.32e+02	216	4	25.0	167	2	S65670	phosphoprotein phosph	9.32e+02
144	4	25.0	73	2	A24836	hypothetical protein	9.32e+02	217	4	25.0	167	2	C71700	hypothetical protein	9.32e+02
145	4	25.0	73	2	F69087	hypothetical protein	9.32e+02	218	4	25.0	169	2	A64396	hypothetical protein	9.32e+02
146	4	25.0	75	2	JC6048	regC protein - Serrat	9.32e+02	219	4	25.0	169	2	S37892	hypothetical protein	9.32e+02
147	4	25.0	77	2	S60828	M protein precursor -	9.32e+02	220	4	25.0	173	2	JC2242	superfast myosin ligh	9.32e+02
148	4	25.0	83	2	C70073	hypothetical protein	9.32e+02	221	4	25.0	173	1	GCBOH	glycine cleavage syst	9.32e+02
149	4	25.0	86	2	S33982	trans-activating tran	9.32e+02	222	4	25.0	174	2	S49245	19K protein - soil-bo	9.32e+02
150	4	25.0	86	2	JC5591	transactivator protei	9.32e+02	223	4	25.0	174	2	H71267	conserved hypothetical	9.32e+02
151	4	25.0	86	2	S54381	tat protein - human i	9.32e+02	224	4	25.0	174	1	DNBPMU	DNA-binding protein -	9.32e+02
152	4	25.0	87	2	T03360	gene e7 protein - Lac	9.32e+02	225	4	25.0	174	2	A64619	polypeptide deformyla	9.32e+02
153	4	25.0	88	1	EBE51	immediate-early-5 pro	9.32e+02	226	4	25.0	175	2	S09784	hypothetical protein	9.32e+02
154	4	25.0	90	2	E71150	hypothetical protein	9.32e+02	227	4	25.0	175	2	T03017	probable chitinase (E	9.32e+02
155	4	25.0	90	2	S74568	hypothetical protein	9.32e+02	228	4	25.0	178	2	G71198	hypothetical protein	9.32e+02
156	4	25.0	91	3	T00167	hypothetical protein	9.32e+02	229	4	25.0	179	2	S52812	hypothetical protein,	9.32e+02
157	4	25.0	91	2	A65082	hypothetical protein	9.32e+02	230	4	25.0	182	2	G64835	ymbA protein - Escher	9.32e+02
158	4	25.0	93	2	A71196	hypothetical protein	9.32e+02	231	4	25.0	182	2	S46667	meiosis-specific prot	9.32e+02
159	4	25.0	95	1	S02185	hypothetical protein	9.32e+02	232	4	25.0	182	2	T01909	transcription initiat	9.32e+02
160	4	25.0	95	1	TNLJ12	trans-activating tran	9.32e+02	233	4	25.0	185	2	JC4667	TB2/DPI protein homol	9.32e+02
161	4	25.0	96	1	THOAA	glycoprotein hormones	9.32e+02	234	4	25.0	185	1	A44559	relaxin 1 precursor -	9.32e+02
162	4	25.0	96	2	A05096	glycoprotein hormones	9.32e+02	235	4	25.0	187	2	H71819	hypothetical protein	9.32e+02
163	4	25.0	98	2	S75283	hypothetical protein	9.32e+02	236	4	25.0	187	2	G64883	ydal protein - Escher	9.32e+02
164	4	25.0	101	2	D71604	small nuclear ribonuc	9.32e+02	237	4	25.0	190	2	S34309	probable transcriptio	9.32e+02
165	4	25.0	102	2	S31417	urease (EC 3.5.1.5) 1	9.32e+02	238	4	25.0	191	2	J01999	hypothetical 21.5K pr	9.32e+02
166	4	25.0	106	2	S80232	gibberellin-regulated	9.32e+02	239	4	25.0	191	2	A39265	GTP-binding protein G	9.32e+02
167	4	25.0	108	2	A60627	glycoprotein hormones	9.32e+02	240	4	25.0	192	2	S70996	espA protein - Escher	9.32e+02
168	4	25.0	111	2	B74041	hypothetical protein	9.32e+02	241	4	25.0	192	1	MOCHLA	myosin alkali light c	9.32e+02
169	4	25.0	112	2	S06563	finger protein (clone	9.32e+02	242	4	25.0	192	2	A55924	cell division control	9.32e+02



243	4	25.0	195	2	G70211	hypothetical protein	9.32e+02	316	4	25.0	249	2	C70436	conserved hypothetical	9.32e+02
244	4	25.0	196	2	S06558	finger protein (clone	9.32e+02	317	4	25.0	250	2	S30427	Mst26Aa protein - fru	9.32e+02
245	4	25.0	197	2	S06581	finger protein (clone	9.32e+02	318	4	25.0	251	2	S74080	arPA protein - Strept	9.32e+02
246	4	25.0	198	2	P00327	heparin-binding fibro	9.32e+02	319	4	25.0	252	2	S73889	probable lipoprotein	9.32e+02
247	4	25.0	199	2	B71114	hypothetical protein	9.32e+02	320	4	25.0	253	2	S48848	chitinase (EC 3.2.1.1	9.32e+02
248	4	25.0	200	2	S01366	spermine-binding prot	9.32e+02	321	4	25.0	254	2	S55288	hypothetical 27.4K pr	9.32e+02
249	4	25.0	201	2	S61090	hypothetical protein	9.32e+02	322	4	25.0	255	2	E64420	uroporphyrin-III C-me	9.32e+02
250	4	25.0	203	2	C70484	peptidoglycan associ	9.32e+02	323	4	25.0	256	2	T03091	hypothetical protein	9.32e+02
251	4	25.0	203	2	I50481	synapsin protein SNAP-	9.32e+02	324	4	25.0	257	2	S70719	hypothetical protein	9.32e+02
252	4	25.0	203	2	S53700	hypothetical protein	9.32e+02	325	4	25.0	258	2	S30429	Mst26Aa protein - fru	9.32e+02
253	4	25.0	204	2	I50480	synapse protein SNAP-	9.32e+02	326	4	25.0	259	1	QRECM3	membrane-bound iron(I	9.32e+02
254	4	25.0	205	2	G70304	hypothetical protein	9.32e+02	327	4	25.0	260	2	I37532	MHC class II histocom	9.32e+02
255	4	25.0	205	2	E64621	conserved hypothetical	9.32e+02	328	4	25.0	261	2	E71601	probable integral mem	9.32e+02
256	4	25.0	206	2	A37861	synaptosomal-associat	9.32e+02	329	4	25.0	262	1	WMTMP3	30K protein - pepper	9.32e+02
257	4	25.0	206	2	A33623	synaptosomal-associat	9.32e+02	330	4	25.0	263	1	WBBPT5	gene D11 protein - ph	9.32e+02
258	4	25.0	206	2	I67823	nerve terminal protei	9.32e+02	331	4	25.0	264	2	G64216	ribosomal protein L3	9.32e+02
259	4	25.0	207	2	C71914	phosphoserine phosph	9.32e+02	332	4	25.0	265	2	S74631	hypothetical protein	9.32e+02
260	4	25.0	207	2	D64601	phosphoserine phosph	9.32e+02	333	4	25.0	266	2	I46080	uropodakin Ia - bovine	9.32e+02
261	4	25.0	208	2	S29081	SKY protein homolog -	9.32e+02	334	4	25.0	267	2	A36949	28.9K basic DNA-bind	9.32e+02
262	4	25.0	209	2	J71322	hypothetical protein	9.32e+02	335	4	25.0	268	2	S10321	31K protein - frog vi	9.32e+02
263	4	25.0	210	2	JC5512	SNARE protein 23 - mo	9.32e+02	336	4	25.0	269	2	E64390	coenzyme F420 hydroge	9.32e+02
264	4	25.0	210	2	I50552	synapse protein - mar	9.32e+02	337	4	25.0	270	2	B70073	glucose 1-dehydrogena	9.32e+02
265	4	25.0	210	2	A64394	hypothetical protein	9.32e+02	338	4	25.0	271	1	EQXL	enkephalin precursor	9.32e+02
266	4	25.0	211	2	JC5296	vesicle-membrane fusi	9.32e+02	339	4	25.0	272	2	S30415	male accessory gland	9.32e+02
267	4	25.0	214	2	D71282	probable GedaA protein	9.32e+02	340	4	25.0	273	2	S30425	Mst26Aa protein - fru	9.32e+02
268	4	25.0	214	1	KIECA	adenylate kinase (EC	9.32e+02	341	4	25.0	274	2	G64548	hypothetical protein	9.32e+02
269	4	25.0	215	1	D22794	phosphoribosylanthran	9.32e+02	342	4	25.0	275	3	JC6515	beta 2 toxin - Clostr	9.32e+02
270	4	25.0	215	2	E42148	GTP-binding protein r	9.32e+02	343	4	25.0	276	2	I49059	Ly49c - mouse	9.32e+02
271	4	25.0	216	2	C70347	hypothetical protein	9.32e+02	344	4	25.0	277	2	I49050	Ly-49E-GE antigen - m	9.32e+02
272	4	25.0	217	2	A54688	modular adaptor Grb2	9.32e+02	345	4	25.0	278	2	I49114	Ly49H - mouse	9.32e+02
273	4	25.0	217	2	G70407	hypothetical protein	9.32e+02	346	4	25.0	279	2	I49363	natural killer cell r	9.32e+02
274	4	25.0	217	2	S26050	growth factor recepto	9.32e+02	347	4	25.0	280	2	A27356	nuclease Sm (EC 3.1.3	9.32e+02
275	4	25.0	217	2	A43321	growth factor recepto	9.32e+02	348	4	25.0	281	2	T00464	hypothetical protein	9.32e+02
276	4	25.0	221	2	S16252	trypsin inhibitor hom	9.32e+02	349	4	25.0	282	2	S75846	hypothetical protein	9.32e+02
277	4	25.0	222	2	S52380	NADH dehydrogenase (u	9.32e+02	350	4	25.0	283	2	JQ1473	pancreatic elastase (	9.32e+02
278	4	25.0	224	2	S06545	finger protein (clone	9.32e+02	351	4	25.0	284	2	S49049	probable S-adenosylme	9.32e+02
279	4	25.0	225	2	S77377	hypothetical protein	9.32e+02	352	4	25.0	285	2	G64691	UDP-N-acetylglucosami	9.32e+02
280	4	25.0	226	2	H1181	probable aspartate ra	9.32e+02	353	4	25.0	286	2	T71615	protein of the YMR7 f	9.32e+02
281	4	25.0	226	2	H71700	hypothetical protein	9.32e+02	354	4	25.0	287	1	NIICLP	nitrogenase (EC 1.18.	9.32e+02
282	4	25.0	226	2	G64387	hypothetical protein	9.32e+02	355	4	25.0	288	2	F64590	diaminopimelate epime	9.32e+02
283	4	25.0	227	2	A48412	hypothetical protein	9.32e+02	356	4	25.0	289	2	S10186	pyrroline-5-carboxyla	9.32e+02
284	4	25.0	228	2	S20416	ASP-56 protein - pig	9.32e+02	357	4	25.0	290	2	S34977	galactosyltransferase	9.32e+02
285	4	25.0	229	2	D64663	tRNA - Helicobacter p	9.32e+02	358	4	25.0	291	2	D71226	hypothetical protein	9.32e+02
286	4	25.0	231	2	B36841	C20L protein - variol	9.32e+02	359	4	25.0	292	2	S57381	protein disulfide iso	9.32e+02
287	4	25.0	231	2	S34448	small nuclear ribonuc	9.32e+02	360	4	25.0	293	2	B40825	hypothetical protein	9.32e+02
288	4	25.0	232	2	A27503	testis-specific prote	9.32e+02	361	4	25.0	294	2	S06572	finger protein (clone	9.32e+02
289	4	25.0	232	2	I83349	CS6 subunits chaperon	9.32e+02	362	4	25.0	295	2	T03165	hypothetical protein	9.32e+02
290	4	25.0	235	2	G71259	probable ABC transpor	9.32e+02	363	4	25.0	296	2	S63225	ribosomal protein yml	9.32e+02
291	4	25.0	237	2	A20688	Balbani ring 2 chain	9.32e+02	364	4	25.0	297	2	T02813	hypothetical protein	9.32e+02
292	4	25.0	237	2	B65135	hypothetical 27.1 kD	9.32e+02	365	4	25.0	298	2	E71145	probable cobalamin bi	9.32e+02
293	4	25.0	238	2	S48693	fluorescent protein -	9.32e+02	366	4	25.0	299	2	JE0235	HIV-promoter GC-rich	9.32e+02
294	4	25.0	238	2	S51331	green fluorescent pro	9.32e+02	367	4	25.0	300	2	JC5661	hepatoma-derived grow	9.32e+02
295	4	25.0	238	2	S51330	green fluorescent pro	9.32e+02	368	4	25.0	301	2	F64338	agmatinase (EC 3.5.3.	9.32e+02
296	4	25.0	238	2	S79334	CotB protein precurs	9.32e+02	369	4	25.0	302	2	A41489	flagellar filament 31	9.32e+02
297	4	25.0	238	2	JQ1514	green-fluorescent pro	9.32e+02	370	4	25.0	303	2	B64105	naphthoate synthase (	9.32e+02
298	4	25.0	239	2	E64699	hypothetical protein	9.32e+02	371	4	25.0	304	2	S76787	hypothetical protein	9.32e+02
299	4	25.0	239	2	S64573	hypothetical protein	9.32e+02	372	4	25.0	305	2	A61267	arylamine N-acetyltra	9.32e+02
300	4	25.0	240	2	H69533	hypothetical protein	9.32e+02	373	4	25.0	306	1	B64113	acetyl-CoA carboxylas	9.32e+02
301	4	25.0	241	2	JE0169	dnau heat shock prote	9.32e+02	374	4	25.0	307	2	B71013	hypothetical protein	9.32e+02
302	4	25.0	241	2	H64692	hypothetical protein	9.32e+02	375	4	25.0	308	2	F69719	succinate--CoA ligase	9.32e+02
303	4	25.0	241	2	C46181	pX-tax-orf II (altern	9.32e+02	376	4	25.0	309	2	G70426	aminomethyltransferas	9.32e+02
304	4	25.0	242	2	JC4300	galectin-3 - rabbit	9.32e+02	377	4	25.0	310	2	P3WMCC	3a protein - cowpea c	9.32e+02
305	4	25.0	243	2	I51746	MHC class II alpha ch	9.32e+02	378	4	25.0	311	2	F64042	GTP-binding protein e	9.32e+02
306	4	25.0	243	2	D64300	(R)-2-hydroxyglutaryl	9.32e+02	379	4	25.0	312	1	GEMSN	osteonectin precursor	9.32e+02
307	4	25.0	243	2	A70103	conserved hypothetical	9.32e+02	380	4	25.0	313	2	S71890	immune-factor 1, type	9.32e+02
308	4	25.0	244	2	B69405	uroporphyrin-III C-me	9.32e+02	381	4	25.0	314	2	A64423	modification methylas	9.32e+02
309	4	25.0	244	2	S70176	yenR protein - Versin	9.32e+02	382	4	25.0	315	2	S47276	prolinase - Lactobaci	9.32e+02
310	4	25.0	244	2	H64347	hypothetical protein	9.32e+02	383	4	25.0	316	2	A64410	ornithine carbamoyltr	9.32e+02
311	4	25.0	246	2	G64147	hypothetical protein	9.32e+02	384	4	25.0	317	2	B64587	cag pathogenicity isl	9.32e+02
312	4	25.0	247	2	G71052	hypothetical protein	9.32e+02	385	4	25.0	318	2	D70657	probable mrr - Mycoba	9.32e+02
313	4	25.0	247	2	H69276	L-isoaspartyl-protein	9.32e+02	386	4	25.0	319	2	T01818	membrane protein yeda	9.32e+02
314	4	25.0	248	2	C40635	indole-3-glycerol-pho	9.32e+02	387	4	25.0	320	2	T01818	hypothetical protein	9.32e+02
315	4	25.0	249	2	S38309	SNAP-25 protein - chi	9.32e+02	388	4	25.0	321	2	I36791	hypothetical protein	9.32e+02

389	4	25.0	309	2	H71328	probable flagellar as	9.32e+02	462	373	2	A70334	hypothetical protein	9.32e+02
390	4	25.0	313	2	A69401	conserved hypothetical	9.32e+02	463	373	2	D64401	hypothetical protein	9.32e+02
391	4	25.0	314	2	A38888	COI intron 8 protein	9.32e+02	464	374	1	A56669	alcohol dehydrogenase	9.32e+02
392	4	25.0	314	2	S22513	ethylene-forming enzy	9.32e+02	465	374	1	A56643	alcohol dehydrogenase	9.32e+02
393	4	25.0	314	2	I50811	MHC class I protein,	9.32e+02	466	375	1	S66272	alcohol dehydrogenase	9.32e+02
394	4	25.0	317	2	T01452	hypothetical protein	9.32e+02	467	375	2	B65115	hypothetical 43.1 kD	9.32e+02
395	4	25.0	317	2	T01222	hypothetical protein	9.32e+02	468	375	1	S29343	alcohol dehydrogenase	9.32e+02
396	4	25.0	317	2	S35282	eaA protein - phage P	9.32e+02	469	375	1	DECHAL	alcohol dehydrogenase	9.32e+02
397	4	25.0	320	2	S73413	transcription antiter	9.32e+02	470	375	1	I60973	alcohol dehydrogenase	9.32e+02
398	4	25.0	323	2	H71481	hypothetical protein	9.32e+02	471	376	2	A64587	cag pathogenicity isl	9.32e+02
399	4	25.0	325	2	I59998	interferon regulatory	9.32e+02	472	376	1	A64668	alcohol dehydrogenase	9.32e+02
400	4	25.0	325	2	B31595	interferon regulatory	9.32e+02	473	376	2	F64705	conserved hypothetical	9.32e+02
401	4	25.0	326	2	B37314	regulatory protein Ic	9.32e+02	474	376	2	G71925	cag island protein -	9.32e+02
402	4	25.0	327	2	S39604	class I histocompatib	9.32e+02	475	377	2	A53044	geranylgeranyl-diphos	9.32e+02
403	4	25.0	328	2	A63330	interferon regulatory	9.32e+02	476	377	2	B53044	geranylgeranyl-diphos	9.32e+02
404	4	25.0	328	2	S70885	UDP-galactose 4-epime	9.32e+02	477	377	2	S27671	rfp protein - Shigell	9.32e+02
405	4	25.0	328	2	F64187	p-aminobenzoate synth	9.32e+02	478	378	2	A55735	G protein-coupled rec	9.32e+02
406	4	25.0	328	2	B54058	UDP-N-acetylglucosam	9.32e+02	479	378	2	S44777	C30A5.3 protein - Cae	9.32e+02
407	4	25.0	329	2	A31595	interferon regulatory	9.32e+02	480	379	1	S71244	alcohol dehydrogenase	9.32e+02
408	4	25.0	330	2	B27246	L-lactate dehydrogena	9.32e+02	481	380	2	S66728	hypothetical protein	9.32e+02
409	4	25.0	331	2	S45584	transcription activat	9.32e+02	482	380	1	S57650	alcohol dehydrogenase	9.32e+02
410	4	25.0	331	1	DEGRHM	L-lactate dehydrogena	9.32e+02	483	382	1	DEPJAI	alcohol dehydrogenase	9.32e+02
411	4	25.0	331	2	B40951	TFII protein - mouse	9.32e+02	484	385	2	A71107	hypothetical protein	9.32e+02
412	4	25.0	332	3	T00175	hypothetical protein	9.32e+02	485	385	2	S68067	major outer membrane	9.32e+02
413	4	25.0	332	1	D31265	37K membrane antigen	9.32e+02	486	385	2	S68068	major outer membrane	9.32e+02
414	4	25.0	333	1	A31568	electron transfer fla	9.32e+02	487	385	2	S68070	major outer membrane	9.32e+02
415	4	25.0	334	2	I51656	L-lactate dehydrogena	9.32e+02	488	386	2	D69040	conserved hypothetical	9.32e+02
416	4	25.0	334	1	A40166	gap junction protein	9.32e+02	489	386	2	S14874	ADP,ATP carrier prote	9.32e+02
417	4	25.0	335	2	E71140	hypothetical protein	9.32e+02	490	387	2	S42773	lota toxin component	9.32e+02
418	4	25.0	337	2	I38500	interferon gamma rece	9.32e+02	491	388	2	T02817	exopolysphatase XP	9.32e+02
419	4	25.0	337	2	S70469	fructose-bisphosphata	9.32e+02	492	388	2	F64147	hypothetical protein	9.32e+02
420	4	25.0	339	2	S39768	neutrophil cytosol fa	9.32e+02	493	389	2	A71091	hypothetical protein	9.32e+02
421	4	25.0	339	2	H71361	conserved hypothetical	9.32e+02	494	392	2	H71520	probable hth transcri	9.32e+02
422	4	25.0	339	2	T08618	PAP protein - Leishm	9.32e+02	495	392	2	S37902	hypothetical protein	9.32e+02
423	4	25.0	341	2	S60431	hypothetical protein	9.32e+02	496	395	1	S38821	homeotic protein lim-	9.32e+02
424	4	25.0	341	2	E70147	basic membrane lipopr	9.32e+02	497	396	2	JD0003	xylose (EC 3.2.1.-)	9.32e+02
425	4	25.0	342	2	S36367	nmtl protein - Asperg	9.32e+02	498	396	2	S34963	rfbX protein - Shigel	9.32e+02
426	4	25.0	342	2	F69261	conserved hypothetical	9.32e+02	499	398	2	S37630	poly(beta-D-mannurona	9.32e+02
427	4	25.0	343	2	T01068	calreticulin - Africa	9.32e+02	500	399	2	G70328	chromosome replicatio	9.32e+02
428	4	25.0	343	2	E70856	probable pfkA protein	9.32e+02	501	399	2	D64015	hypothetical protein	9.32e+02
429	4	25.0	344	2	S50474	hypothetical protein	9.32e+02	502	400	2	S43376	calreticulin, brain i	9.32e+02
430	4	25.0	344	2	JN0055	Pilr protein - Pseudo	9.32e+02	503	401	2	S59670	RAD17 protein - yeast	9.32e+02
431	4	25.0	346	2	T71042	probable threonine sy	9.32e+02	504	402	2	D70186	xylose operon regulat	9.32e+02
432	4	25.0	346	2	I64237	hypothetical protein	9.32e+02	505	402	2	T00623	hypothetical protein	9.32e+02
433	4	25.0	346	1	WMV22W	3beta-hydroxy-Delta5-	9.32e+02	506	402	2	S32142	hypothetical protein	9.32e+02
434	4	25.0	347	2	S55629	hypothetical protein	9.32e+02	507	403	2	G70043	transporter homolog y	9.32e+02
435	4	25.0	348	2	I37271	cylicin II - human	9.32e+02	508	404	2	T03110	hypothetical protein	9.32e+02
436	4	25.0	348	2	S23398	rhodopsin - Chinese h	9.32e+02	509	405	2	B70300	translation elongatio	9.32e+02
437	4	25.0	348	2	S44628	f22b7.1 protein - Cae	9.32e+02	510	406	2	A56637	calreticulin homolog	9.32e+02
438	4	25.0	348	1	OOHU	rhodopsin - human	9.32e+02	511	407	2	E71374	probable hemolysin -	9.32e+02
439	4	25.0	348	2	S51677	rhodopsin - rat	9.32e+02	512	409	2	D64216	hypothetical protein	9.32e+02
440	4	25.0	354	2	S40508	G-protein alpha subun	9.32e+02	513	410	2	I50494	serine proteinase inh	9.32e+02
441	4	25.0	355	1	DCSCP	phosphoribosylaminoim	9.32e+02	514	412	2	S62239	hypothetical protein	9.32e+02
442	4	25.0	356	2	S42096	interleukin-8 recepto	9.32e+02	515	414	2	H70219	hypothetical protein	9.32e+02
443	4	25.0	356	2	S61766	protein kinase ATN1 (	9.32e+02	516	414	2	S74699	ubiH protein - Synech	9.32e+02
444	4	25.0	357	2	E71708	hypothetical protein	9.32e+02	517	414	2	A701646	5-aminolevulinic acid	9.32e+02
445	4	25.0	359	2	G69822	conserved hypothetical	9.32e+02	518	416	2	JH0819	calreticulin precurs	9.32e+02
446	4	25.0	359	2	E71960	probable peptide meth	9.32e+02	519	416	2	S06763	calreticulin precurs	9.32e+02
447	4	25.0	359	1	ADSC2A	fructose-bisphosphate	9.32e+02	520	417	2	B40016	matrin 3 - human (fra	9.32e+02
448	4	25.0	360	2	S43611	interleukin-8 recepto	9.32e+02	521	417	2	G44417	hypothetical protein	9.32e+02
449	4	25.0	361	2	PN0020	fibroblast growth fac	9.32e+02	522	417	2	A37047	calreticulin precurs	9.32e+02
450	4	25.0	363	2	JC5536	C-Fringe protein 1 pr	9.32e+02	523	419	2	S71343	calreticulin precurs	9.32e+02
451	4	25.0	363	2	S76156	hypothetical protein	9.32e+02	524	419	1	K8XL2B	keratin, 64K type II	9.32e+02
452	4	25.0	364	2	C71295	probable octaprenyl-d	9.32e+02	525	421	2	S67999	calreticulin precurs	9.32e+02
453	4	25.0	365	2	A48135	transcription negativ	9.32e+02	526	421	2	A24810	phaseolin beta chain	9.32e+02
454	4	25.0	365	2	T70043	hypothetical protein	9.32e+02	527	422	2	JE0239	lin-10 protein - rat	9.32e+02
455	4	25.0	365	2	JH0543	class I histocompatib	9.32e+02	528	422	1	BMRT2Y	synaptotagmin II - ra	9.32e+02
456	4	25.0	367	2	C41102	di-N-acetylchitinobias	9.32e+02	529	423	2	S43886	DNA methyltransferase	9.32e+02
457	4	25.0	367	2	EJ0104	testicular serine pro	9.32e+02	530	425	2	C70450	conserved hypothetical	9.32e+02
458	4	25.0	367	2	E70347	cell division protein	9.32e+02	531	425	2	S53004	mitosis-specific cycl	9.32e+02
459	4	25.0	369	2	B71624	hypothetical protein	9.32e+02	532	425	2	JU0467	3-deoxy-D-manno-ocul	9.32e+02
460	4	25.0	369	2	S74017	hypothetical protein	9.32e+02	533	426	2	D42519	A20R protein - vaccin	9.32e+02
461	4	25.0	373	2	S71218	gene At103 protein -	9.32e+02	534	426	2	D64970	hypothetical protein	9.32e+02

535	4	25.0	427	2	I49603	transcription regulat	9.32e+02	608	4	25.0	496	2	A49930	carB protein homolog	9.32e+02
536	4	25.0	428	2	A42971	cytochrome P450terp -	9.32e+02	609	4	25.0	498	2	C36851	A29L protein - variol	9.32e+02
537	4	25.0	429	2	A47305	translation initiatio	9.32e+02	610	4	25.0	499	2	I48741	ryanodine receptor ty	9.32e+02
538	4	25.0	429	2	S24991	glucose-1-phosphate a	9.32e+02	611	4	25.0	499	2	F69001	hypothetical protein	9.32e+02
539	4	25.0	431	2	E71524	probable aspartokinas	9.32e+02	612	4	25.0	500	1	ISEBAB	L-arabinose isomerase	9.32e+02
540	4	25.0	434	2	F71638	UDP-glucose 6-dehydro	9.32e+02	613	4	25.0	500	2	G69428	hypothetical protein	9.32e+02
541	4	25.0	436	2	T00756	hypothetical protein	9.32e+02	614	4	25.0	502	2	D65057	hypothetical protein	9.32e+02
542	4	25.0	436	2	A23498	phaseolin type alpha	9.32e+02	615	4	25.0	502	1	ACHUA7	nicotinic acetylcholi	9.32e+02
543	4	25.0	437	2	A31752	transcription factor	9.32e+02	616	4	25.0	505	2	S64837	hypothetical protein	9.32e+02
544	4	25.0	437	2	S42111	transcription factor	9.32e+02	617	4	25.0	507	1	PW2MA	H+-transporting ATP s	9.32e+02
545	4	25.0	438	2	G02822	VIP2 receptor - human	9.32e+02	618	4	25.0	508	2	A35865	hydroxymethylglutaryl	9.32e+02
546	4	25.0	438	2	JC2463	vasoactive intestinal	9.32e+02	619	4	25.0	509	1	PKWTAM	H+-transporting ATP s	9.32e+02
547	4	25.0	438	2	H71138	probable phosphoribos	9.32e+02	620	4	25.0	510	1	FOLJTM	gag polyprotein - sim	9.32e+02
548	4	25.0	439	2	E71116	hypothetical protein	9.32e+02	621	4	25.0	511	2	D64905	probable sugar transp	9.32e+02
549	4	25.0	440	2	B70357	hypothetical protein	9.32e+02	622	4	25.0	515	3	JE0361	cytochromes P450, CYP	9.32e+02
550	4	25.0	440	2	JE0085	PI protein - Filament	9.32e+02	623	4	25.0	518	2	H69649	2-isopropylmalate syn	9.32e+02
551	4	25.0	445	4	A28516	probable cytochrome P	9.32e+02	624	4	25.0	521	2	F71009	probable sodium-depen	9.32e+02
552	4	25.0	447	2	D71611	hypothetical protein	9.32e+02	625	4	25.0	522	2	G02533	occludin - human	9.32e+02
553	4	25.0	448	2	E56695	transducin-like enhan	9.32e+02	626	4	25.0	522	2	S52491	polyadenylate binding	9.32e+02
554	4	25.0	450	2	C71474	probable muramoyl-dap	9.32e+02	627	4	25.0	524	2	E49600	probable aphid transm	9.32e+02
555	4	25.0	450	2	A27053	lipoprotein lipase (E	9.32e+02	628	4	25.0	524	1	VGWNB8	spike glycoprotein pr	9.32e+02
556	4	25.0	450	2	A27053	lipoprotein lipase (E	9.32e+02	629	4	25.0	524	2	S55097	probable membrane pr	9.32e+02
557	4	25.0	453	1	IQXMC	replication initiatio	9.32e+02	630	4	25.0	525	2	C69794	glutamate synthase (f	9.32e+02
558	4	25.0	453	2	S46700	SSF1 protein - yeast	9.32e+02	631	4	25.0	526	2	S35190	clostripain (EC 3.4.2	9.32e+02
559	4	25.0	454	2	B41621	env polyprotein D - h	9.32e+02	632	4	25.0	526	2	I49134	prolyl 4-hydroxylase	9.32e+02
560	4	25.0	455	2	S68693	hexokinase (EC 2.7.1.	9.32e+02	633	4	25.0	529	1	SAHU4F	cell surface antigen	9.32e+02
561	4	25.0	456	2	JC5521	TATA-binding protein-	9.32e+02	634	4	25.0	529	2	A43505	listeriolysin O precu	9.32e+02
562	4	25.0	457	2	A55199	odr-7 protein - Caeno	9.32e+02	635	4	25.0	530	2	A45690	transactivator EBNA-2	9.32e+02
563	4	25.0	460	2	F71292	probable GTP-binding	9.32e+02	636	4	25.0	531	2	S30011	probable membrane pr	9.32e+02
564	4	25.0	460	2	S76894	hypothetical protein	9.32e+02	637	4	25.0	533	2	A56110	tyrosine phosphoprote	9.32e+02
565	4	25.0	462	2	E61019	hypothetical protein	9.32e+02	638	4	25.0	533	2	S52702	REF2 protein - yeast	9.32e+02
566	4	25.0	462	2	S10397	finger protein koxl -	9.32e+02	639	4	25.0	535	2	S74703	hypothetical protein	9.32e+02
567	4	25.0	462	2	G70753	probable amidase - My	9.32e+02	640	4	25.0	536	2	S06548	finger protein (clone	9.32e+02
568	4	25.0	462	2	A64416	adenylosuccinate lyas	9.32e+02	641	4	25.0	537	2	S35519	hypothetical protein	9.32e+02
569	4	25.0	464	2	A47655	spliceosome-associat	9.32e+02	642	4	25.0	538	2	F71255	conserved hypothetica	9.32e+02
570	4	25.0	465	2	S46759	hypothetical protein	9.32e+02	643	4	25.0	538	1	VGNZTR	cell fusion glycoprot	9.32e+02
571	4	25.0	465	2	S68462	protein kinase ARPK6/	9.32e+02	644	4	25.0	540	2	JC4916	signal transducing ad	9.32e+02
572	4	25.0	467	2	B70116	transhydrogenase chal	9.32e+02	645	4	25.0	540	2	JC4917	site-specific DNA-met	9.32e+02
573	4	25.0	467	2	H65018	ethanolamine ammonia-	9.32e+02	646	4	25.0	540	2	JC4917	metalloproteinase (EC	9.32e+02
574	4	25.0	469	2	E64319	carbon-monoxide dehyd	9.32e+02	647	4	25.0	543	2	B54424	dipeptide ABC transpo	9.32e+02
575	4	25.0	471	2	A41680	integral membrane pro	9.32e+02	648	4	25.0	543	2	A54116	actosomal protein sp3	9.32e+02
576	4	25.0	471	2	B57229	hydroxymethylglutaryl	9.32e+02	649	4	25.0	544	2	S42723	cytochrome P450 1B1 -	9.32e+02
577	4	25.0	472	2	E70325	hypothetical protein	9.32e+02	650	4	25.0	545	2	S46151	matricin - mouse	9.32e+02
578	4	25.0	472	2	A69018	methyl viologen-reduc	9.32e+02	651	4	25.0	548	2	S46151	probable purine nucle	9.32e+02
579	4	25.0	474	2	JH0790	lipoprotein lipase (E	9.32e+02	652	4	25.0	548	2	JC4917	signal transducing ad	9.32e+02
580	4	25.0	475	2	S31927	drop9 protein - human	9.32e+02	653	4	25.0	549	2	S48169	metalloproteinase (EC	9.32e+02
581	4	25.0	476	2	B44997	merozoite surface ant	9.32e+02	654	4	25.0	551	2	G69618	dipeptide ABC transpo	9.32e+02
582	4	25.0	477	2	S19679	1-aminocyclopropane-1	9.32e+02	655	4	25.0	551	2	S30691	multi-drug resistance	9.32e+02
583	4	25.0	477	2	A42925	hyaluronan receptor R	9.32e+02	656	4	25.0	552	2	JQ1857	arylsulfatase (EC 3.1	9.32e+02
584	4	25.0	477	2	S51344	probable membrane pro	9.32e+02	657	4	25.0	552	2	JQ1857	DNA ligase (ATP) (EC	9.32e+02
585	4	25.0	477	2	T01828	hypothetical protein	9.32e+02	658	4	25.0	555	2	D71714	ABC transporter ATP-b	9.32e+02
586	4	25.0	478	2	S51457	toxin apxIII secretio	9.32e+02	659	4	25.0	556	2	A53376	tryptophan monooxygen	9.32e+02
587	4	25.0	478	2	A32555	probable membrane pro	9.32e+02	660	4	25.0	559	1	C9HU	complement C9 precurs	9.32e+02
588	4	25.0	480	2	B64308	major merozoite surfa	9.32e+02	661	4	25.0	559	2	S55383	peptidylprolyl isomer	9.32e+02
589	4	25.0	481	2	JE0132	hypothetical protein	9.32e+02	662	4	25.0	561	2	S71189	Dwarfl protein - Arab	9.32e+02
590	4	25.0	481	2	E64623	glucose-1-phosphate a	9.32e+02	663	4	25.0	562	2	S56145	BS99 protein - human	9.32e+02
591	4	25.0	482	2	A44997	IMP dehydrogenase (EC	9.32e+02	664	4	25.0	563	2	C64420	N-methylhydantoinase	9.32e+02
592	4	25.0	482	2	S22654	merozoite surface ant	9.32e+02	665	4	25.0	564	2	I61767	keratin type II - hum	9.32e+02
593	4	25.0	484	2	A58663	ARS-binding factor-1	9.32e+02	666	4	25.0	566	2	S75233	ABC-type transport pr	9.32e+02
594	4	25.0	484	2	A58663	catalase (EC 1.11.1.6	9.32e+02	667	4	25.0	568	2	A55377	CPE-binding protein -	9.32e+02
595	4	25.0	487	1	LQBP34	DNA ligase (ATP) (EC	9.32e+02	668	4	25.0	569	2	C69471	probable fatty-acid--	9.32e+02
596	4	25.0	487	1	AJSMRC	argininosuccinate syn	9.32e+02	669	4	25.0	570	2	A45624	trophozoite cysteine	9.32e+02
597	4	25.0	487	2	JT0407	keratin 8, type II cy	9.32e+02	670	4	25.0	571	2	I59402	NABL - rat	9.32e+02
598	4	25.0	488	1	OXASBI	mRNA maturase bil - E	9.32e+02	671	4	25.0	571	2	S00566	convicilin precursor	9.32e+02
599	4	25.0	489	2	B53153	glucose transport pro	9.32e+02	672	4	25.0	573	2	B71324	probable DNA repair p	9.32e+02
600	4	25.0	490	2	A25954	cytochrome P450 2C6,	9.32e+02	673	4	25.0	575	2	I40615	flagellin protein - C	9.32e+02
601	4	25.0	491	2	S58202	HMGs protein - yeast	9.32e+02	674	4	25.0	575	2	S41311	flagellin B - Campylo	9.32e+02
602	4	25.0	492	2	S63065	hypothetical protein	9.32e+02	675	4	25.0	576	2	S15286	flagellin - Campyloba	9.32e+02
603	4	25.0	493	2	T02458	hypothetical protein	9.32e+02	676	4	25.0	576	2	B39228	flagellin B - Campylo	9.32e+02
604	4	25.0	493	2	A32454	X-Pro dipeptidase (EC	9.32e+02	677	4	25.0	577	2	E64328	flagellin--trNA ligase	9.32e+02
605	4	25.0	493	1	S72196	X-Pro dipeptidase (EC	9.32e+02	678	4	25.0	579	2	B45266	MPL-K protein precurs	9.32e+02
606	4	25.0	494	1	UF6CDW	aspartate ammonia-lya	9.32e+02	679	4	25.0	580	2	S63391	probable membrane pro	9.32e+02
607	4	25.0	494	1	JQ2156	nitrogenase (EC 1.18.	9.32e+02	680	4	25.0	581	1	KRNS2	keratin, type II cyto	9.32e+02
			496	2	B64591	aminopeptidase a/i -	9.32e+02				583	2	JC5292	GTP-binding protein G	9.32e+02

681	4	25.0	583	2	S56680	9.32e+02	beta-fructofuranosida	754	4	25.0	719	2	A30047	enhancer of split pro	9.32e+02
682	4	25.0	584	2	JC5291	9.32e+02	GTP-binding protein G	755	4	25.0	721	2	S31824	gene Mx protein - duc	9.32e+02
683	4	25.0	586	3	JC5500	9.32e+02	hepatocyte nuclear fa	756	4	25.0	722	2	B61231	myosin heavy chain, n	9.32e+02
684	4	25.0	588	2	I37202	9.32e+02	B-CAM protein - human	757	4	25.0	725	2	D71607	V5F45-like protein (S	9.32e+02
685	4	25.0	589	2	S06954	9.32e+02	intermediate filament	758	4	25.0	722	2	A57148	outer membrane protei	9.32e+02
686	4	25.0	590	2	S66956	9.32e+02	hypothetical protein	759	4	25.0	727	1	VCPVIF	coat protein VPI - fe	9.32e+02
687	4	25.0	591	2	S73708	9.32e+02	MG321 homolog H08_orf	760	4	25.0	727	1	VCPVFP	coat protein VPI - fe	9.32e+02
688	4	25.0	592	2	S56681	9.32e+02	beta-fructofuranosida	761	4	25.0	728	2	I56413	sodium-dependent neur	9.32e+02
689	4	25.0	593	1	K12PMN	9.32e+02	protein kinase cdrl (	762	4	25.0	730	2	I52632	sodium-dependent neur	9.32e+02
690	4	25.0	597	2	A33668	9.32e+02	sterol esterase (EC 3	763	4	25.0	730	2	A36226	collagen alpha 1 chal	9.32e+02
691	4	25.0	599	2	A77928	9.32e+02	signal recognition pa	764	4	25.0	730	2	A26391	phospholipase C (EC 3	9.32e+02
692	4	25.0	600	2	S28496	9.32e+02	microtubule-associate	765	4	25.0	733	2	S23468	ococyte-specific prote	9.32e+02
693	4	25.0	601	2	S33377	9.32e+02	p63 protein - human	766	4	25.0	736	2	G01522	acidic 82 kba protein	9.32e+02
694	4	25.0	601	2	T02581	9.32e+02	hypothetical protein	767	4	25.0	737	1	VCPVCD	coat protein VPI - ca	9.32e+02
695	4	25.0	602	2	JC5906	9.32e+02	major capsid protein	768	4	25.0	738	2	A40096	platelet-endothelial	9.32e+02
696	4	25.0	605	2	G71881	9.32e+02	probable flagellar ba	769	4	25.0	738	2	A53542	brefeldin A-sensitive	9.32e+02
697	4	25.0	605	2	S20007	9.32e+02	beta-conglycinin alph	770	4	25.0	738	2	I40719	isocitrate dehydrogen	9.32e+02
698	4	25.0	606	2	G4659	9.32e+02	flagellar hook-associ	771	4	25.0	739	2	A54808	diastrophic dysplasia	9.32e+02
699	4	25.0	606	2	G71853	9.32e+02	flagellar hook-associ	772	4	25.0	740	2	A47134	alpha-1,3(6)-mannosyl	9.32e+02
700	4	25.0	608	2	S32823	9.32e+02	sonatotropin receptor	773	4	25.0	741	2	JC2074	alpha-1,3(6)-mannosyl	9.32e+02
701	4	25.0	608	2	A53195	9.32e+02	alpha-albumin - rat	774	4	25.0	741	2	B49555	enhancer of split hom	9.32e+02
702	4	25.0	609	2	S65208	9.32e+02	probable membrane pro	775	4	25.0	742	2	A49341	isocitrate dehydrogen	9.32e+02
703	4	25.0	611	1	EBESM	9.32e+02	immediate-early prote	776	4	25.0	742	2	C71983	ATP-dependent protein	9.32e+02
704	4	25.0	618	2	S59067	9.32e+02	probable glycine--tRN	777	4	25.0	743	2	C56695	transducin-like enhan	9.32e+02
705	4	25.0	621	2	S55503	9.32e+02	protein p66 precursor	778	4	25.0	745	2	T03119	hypothetical protein	9.32e+02
706	4	25.0	621	2	S70121	9.32e+02	signal recognition pa	779	4	25.0	754	2	S37403	transcription factor	9.32e+02
707	4	25.0	625	2	S38965	9.32e+02	mannosyl-oligosacchar	780	4	25.0	755	2	A44315	cartilage oligomeric	9.32e+02
708	4	25.0	628	2	I38000	9.32e+02	Lutheran blood group	781	4	25.0	759	2	C64345	cell division control	9.32e+02
709	4	25.0	629	2	A29666	9.32e+02	keratin, 65K type II	782	4	25.0	761	2	T03719	probable thyroid rece	9.32e+02
710	4	25.0	629	2	S42629	9.32e+02	keratin K3 - rabbit	783	4	25.0	765	2	G64502	hypothetical protein	9.32e+02
711	4	25.0	630	2	S70909	9.32e+02	transferrin-binding p	784	4	25.0	770	2	G02228	DOC-2 - human	9.32e+02
712	4	25.0	631	2	JC4298	9.32e+02	hyaluronan receptor -	785	4	25.0	770	2	B56695	transducin-like enhan	9.32e+02
713	4	25.0	631	2	S70910	9.32e+02	transferrin-binding p	786	4	25.0	771	2	S35661	ESG protein - mouse	9.32e+02
714	4	25.0	633	2	E31265	9.32e+02	ipaA protein - Shigel	787	4	25.0	772	2	D56695	transducin-like enhan	9.32e+02
715	4	25.0	633	2	I39585	9.32e+02	dnak-type molecular c	788	4	25.0	772	2	F64052	formate C-acetyltrans	9.32e+02
716	4	25.0	635	2	A45266	9.32e+02	MPI-P protein precurs	789	4	25.0	773	2	S46011	probable pre-mRNA-spl	9.32e+02
717	4	25.0	636	2	I48718	9.32e+02	poly(A) binding prote	790	4	25.0	776	2	S13470	isoamylase (EC 3.2.1.	9.32e+02
718	4	25.0	637	2	A47042	9.32e+02	dnak-type molecular c	791	4	25.0	776	2	A37035	isoamylase (EC 3.2.1.	9.32e+02
719	4	25.0	641	2	A54407	9.32e+02	alpha-mannosidase (EC	792	4	25.0	778	2	I38487	tastin - human	9.32e+02
720	4	25.0	644	2	G70794	9.32e+02	probable alpha-isopro	793	4	25.0	783	2	JC6136	kekin-like protein co	9.32e+02
721	4	25.0	644	2	S53795	9.32e+02	neurofilament protein	794	4	25.0	784	2	A36133	transcription repress	9.32e+02
722	4	25.0	645	2	S20138	9.32e+02	probable protein kina	795	4	25.0	785	2	G64706	subtilisin-like propr	9.32e+02
723	4	25.0	646	2	B70396	9.32e+02	histidine kinase sens	796	4	25.0	813	2	T02672	hypothetical protein	9.32e+02
724	4	25.0	646	2	H71378	9.32e+02	conserved hypothetical	797	4	25.0	814	2	S19658	microbial collagenase	9.32e+02
725	4	25.0	648	2	S73399	9.32e+02	transketolase (EC 2.2	798	4	25.0	817	2	T01866	hypothetical protein	9.32e+02
726	4	25.0	652	2	S11735	9.32e+02	resistance protein Mx	799	4	25.0	821	2	S39983	eps8 protein - mouse	9.32e+02
727	4	25.0	653	2	C70742	9.32e+02	probable rsbu protein	800	4	25.0	821	2	B71229	probable pyruvate, wat	9.32e+02
728	4	25.0	655	2	A54408	9.32e+02	mannosyl-oligosacchar	801	4	25.0	822	2	JT0968	1,4-alpha-glucan bran	9.32e+02
729	4	25.0	656	1	QBET2	9.32e+02	UL25 protein - human	802	4	25.0	823	2	S18968	cyristetin precursor	9.32e+02
730	4	25.0	657	2	G71602	9.32e+02	protein with DnaJ dom	803	4	25.0	823	2	H71300	probable DNA ligase (	9.32e+02
731	4	25.0	659	2	S78554	9.32e+02	mannosyl-oligosacchar	804	4	25.0	823	2	D70386	long-chain-fatty-acid	9.32e+02
732	4	25.0	659	2	S77658	9.32e+02	hypothetical protein	805	4	25.0	824	2	I49584	polycystic kidney dis	9.32e+02
733	4	25.0	660	2	S71276	9.32e+02	beta-fructofuranosida	806	4	25.0	825	2	S75173	hypothetical protein	9.32e+02
734	4	25.0	662	2	A33481	9.32e+02	interferon-induced vi	807	4	25.0	833	1	A31593	heat shock transcript	9.32e+02
735	4	25.0	666	2	S61461	9.32e+02	p83/100 protein - Bor	808	4	25.0	834	2	S51006	probable pyruvate, wat	9.32e+02
736	4	25.0	668	2	I39902	9.32e+02	penicillin-binding pr	809	4	25.0	837	2	A57542	p36 protein - mouse	9.32e+02
737	4	25.0	670	2	S22293	9.32e+02	zinc finger protein A	810	4	25.0	841	2	S69563	suppressor protein PS	9.32e+02
738	4	25.0	670	2	A55001	9.32e+02	catalase (EC 1.11.1.6	811	4	25.0	842	2	H71269	probable translation	9.32e+02
739	4	25.0	671	1	QBRSN	9.32e+02	two-component sensor	812	4	25.0	845	2	S34027	CDC47 protein - yeast	9.32e+02
740	4	25.0	672	2	S61463	9.32e+02	p83/100 protein - Lym	813	4	25.0	848	2	T00788	protein kinase homolo	9.32e+02
741	4	25.0	672	2	C69120	9.32e+02	conserved hypothetical	814	4	25.0	849	2	S00030	neurofilament triplet	9.32e+02
742	4	25.0	679	2	S61460	9.32e+02	p83/100 protein - Lym	815	4	25.0	853	2	D70304	hypothetical protein	9.32e+02
743	4	25.0	684	2	S02666	9.32e+02	novel antigen recepto	816	4	25.0	858	2	T00258	hypothetical protein	9.32e+02
744	4	25.0	686	2	C64288	9.32e+02	hypothetical protein	817	4	25.0	859	1	S06418	3',5'-cyclic-GMP phos	9.32e+02
745	4	25.0	690	2	A47298	9.32e+02	NADPH--ferrithemoprote	818	4	25.0	859	2	S46176	sulfate transport pro	9.32e+02
746	4	25.0	692	2	S71159	9.32e+02	NADPH--ferrithemoprote	819	4	25.0	869	2	A25945	coagulation factor VI	9.32e+02
747	4	25.0	696	2	S71889	9.32e+02	immune factor 1, type	820	4	25.0	870	2	B40121	grape-activating pro	9.32e+02
748	4	25.0	703	2	A64351	9.32e+02	hypothetical protein	821	4	25.0	874	2	JC4930	S-layer protein precu	9.32e+02
749	4	25.0	707	2	H71375	9.32e+02	probable 76K protein	822	4	25.0	881	2	T01269	probable protein kina	9.32e+02
750	4	25.0	708	2	JC6329	9.32e+02	yeast secretory prote	823	4	25.0	885	2	B70393	hypothetical protein	9.32e+02
751	4	25.0	709	1	TWBYS5	9.32e+02	transcription factor	824	4	25.0	886	2	B40585	DNA topoisomerase (AT	9.32e+02
752	4	25.0	713	2	G64606	9.32e+02	hydantoin utilization	825	4	25.0	892	2	S17546	probable serine prote	9.32e+02
753	4	25.0	716	1	A44259	9.32e+02	kinesin-related prote	826	4	25.0	911	2	S70958	otnA protein - Vibrio	9.32e+02

827	4	25.0	912	2	C71004	probable ATP-dependen	9.32e+02	900	4	25.0	1233	2	I54383	gene SBL 8/DXS423E pr	9.32e+02
828	4	25.0	925	1	HXAD41	hexon protein - human	9.32e+02	901	4	25.0	1234	2	T00363	hypothetical protein	9.32e+02
829	4	25.0	928	2	JC2486	3',5'-cyclic-nucleoti	9.32e+02	902	4	25.0	1239	2	G02750	DNA-directed DNA poly	9.32e+02
830	4	25.0	932	2	I52527	PAC54A - mouse (fragm	9.32e+02	903	4	25.0	1259	4	GNHUL1	retrovirus-related re	9.32e+02
831	4	25.0	937	2	I53282	gene PACE4 protein -	9.32e+02	904	4	25.0	1274	2	S28279	hypothetical protein	9.32e+02
832	4	25.0	938	2	B30111	alpha-adaptin C - mou	9.32e+02	905	4	25.0	1281	2	T00346	hypothetical protein	9.32e+02
833	4	25.0	938	2	S11276	alpha-adaptin c - rat	9.32e+02	906	4	25.0	1286	2	S38058	hypothetical protein	9.32e+02
834	4	25.0	947	2	S71922	preprotein translocas	9.32e+02	907	4	25.0	1317	2	B41950	retrovirus-related by	9.32e+02
835	4	25.0	957	2	S66755	probable membrane pro	9.32e+02	908	4	25.0	1322	2	A39612	suppressor of sable p	9.32e+02
836	4	25.0	958	2	S73012	polyketide synthase p	9.32e+02	909	4	25.0	1333	2	S65812	DNA-directed DNA poly	9.32e+02
837	4	25.0	961	1	TSHPU4	thrombospondin 4 prec	9.32e+02	910	4	25.0	1377	2	A38926	DNA-binding protein c	9.32e+02
838	4	25.0	962	2	S60225	ionotropic glutamate	9.32e+02	911	4	25.0	1383	2	A36080	insulin receptor prec	9.32e+02
839	4	25.0	962	2	JC5571	subtilisin-like prote	9.32e+02	912	4	25.0	1400	2	B70963	hypothetical protein	9.32e+02
840	4	25.0	964	2	S25855	DNA-directed DNA poly	9.32e+02	913	4	25.0	1403	2	JQ1397	pros protein - fruit	9.32e+02
841	4	25.0	964	2	S57379	MSH2 protein - yeast	9.32e+02	914	4	25.0	1407	2	A41089	neural precursor pr	9.32e+02
842	4	25.0	969	2	A39490	serine proteinase (EC	9.32e+02	915	4	25.0	1414	2	S23809	collagen alpha 2(I) c	9.32e+02
843	4	25.0	972	2	A30363	glycoprotein GP330, r	9.32e+02	916	4	25.0	1417	2	T03443	receptor-associated c	9.32e+02
844	4	25.0	975	2	I59422	rsec8 - rat (fragment	9.32e+02	917	4	25.0	1420	2	T03749	probable nuclear rece	9.32e+02
845	4	25.0	977	2	A30111	alpha-adaptin A - mou	9.32e+02	918	4	25.0	1433	1	GNVUBW	M polyprotein precurs	9.32e+02
846	4	25.0	985	2	JC4217	alpha-glucosidase (EC	9.32e+02	919	4	25.0	1449	2	D71551	hypothetical protein	9.32e+02
847	4	25.0	988	1	DJVFPP	DNA-directed DNA poly	9.32e+02	920	4	25.0	1458	2	S36014	dynein heavy chain, c	9.32e+02
848	4	25.0	997	2	F71365	probable DNA polymera	9.32e+02	921	4	25.0	1460	2	T00095	hypothetical protein	9.32e+02
849	4	25.0	1001	2	S47243	starch phosphorylase	9.32e+02	922	4	25.0	1465	2	S45628	DNA-directed DNA poly	9.32e+02
850	4	25.0	1001	2	H64593	type III restriction	9.32e+02	923	4	25.0	1473	2	A20872	ovostatin precursor -	9.32e+02
851	4	25.0	1002	2	S64145	sodium-chloride trans	9.32e+02	924	4	25.0	1490	2	JC5145	DNA (cytosine-5'-)met	9.32e+02
852	4	25.0	1004	2	B69483	hypothetical protein	9.32e+02	925	4	25.0	1495	2	S60255	transcription co-repr	9.32e+02
853	4	25.0	1004	2	T00795	tRNA-processing prote	9.32e+02	926	4	25.0	1498	2	S35577	TyB protein - yeast (	9.32e+02
854	4	25.0	1005	2	S19989	hypothetical protein	9.32e+02	927	4	25.0	1501	1	B29813	174K ninac protein -	9.32e+02
855	4	25.0	1005	2	A64465	hypothetical protein	9.32e+02	928	4	25.0	1516	2	E71619	RAD2 endonuclease PFB	9.32e+02
856	4	25.0	1020	2	B39521	glycine dehydrogenase	9.32e+02	929	4	25.0	1547	2	S69842	TyB protein - yeast (	9.32e+02
857	4	25.0	1020	2	JN0124	glycine dehydrogenase	9.32e+02	930	4	25.0	1548	2	S34583	serine proteinase (EC	9.32e+02
858	4	25.0	1021	2	S44644	F37A4.7 protein - Cae	9.32e+02	931	4	25.0	1572	2	T00027	brain-specific angiot	9.32e+02
859	4	25.0	1026	2	T03108	DNA-directed DNA poly	9.32e+02	932	4	25.0	1618	2	T00124	nestin - human	9.32e+02
860	4	25.0	1031	2	A29839	RAD2 protein - yeast	9.32e+02	933	4	25.0	1650	2	S33457	dominant autoantigen	9.32e+02
861	4	25.0	1038	2	JT0663	ras GTPase-activating	9.32e+02	934	4	25.0	1653	2	G65028	hypothetical protein	9.32e+02
862	4	25.0	1040	2	A57638	receptor tyrosine kin	9.32e+02	935	4	25.0	1684	2	T02367	hypothetical protein	9.32e+02
863	4	25.0	1041	2	S42509	Rag-1 protein - chick	9.32e+02	936	4	25.0	1696	2	T00057	hypothetical protein	9.32e+02
864	4	25.0	1044	2	S01966	GTPase-activating pro	9.32e+02	937	4	25.0	1723	2	S58880	receptor DEC-205 - mo	9.32e+02
865	4	25.0	1045	2	S23570	pol polyprotein homol	9.32e+02	938	4	25.0	1729	2	S57596	ribosomal RNA process	9.32e+02
866	4	25.0	1047	2	A40121	GTPase-activating pro	9.32e+02	939	4	25.0	1766	2	S03701	141K protein - pea ea	9.32e+02
867	4	25.0	1054	2	D70425	conserved hypothetical	9.32e+02	940	4	25.0	1772	2	A45532	major merozoite surfa	9.32e+02
868	4	25.0	1062	2	S61196	SUM1 protein - yeast	9.32e+02	941	4	25.0	1777	3	T00490	nonstructural protein	9.32e+02
869	4	25.0	1068	2	S73091	hypothetical protein	9.32e+02	942	4	25.0	1785	2	A45546	major merozoite surfa	9.32e+02
870	4	25.0	1084	2	S57681	CRM1 protein - yeast	9.32e+02	943	4	25.0	1791	2	T02345	hypothetical protein	9.32e+02
871	4	25.0	1091	2	E71322	probable isoleucyl-tr	9.32e+02	944	4	25.0	1802	2	T71616	hypothetical protein	9.32e+02
872	4	25.0	1103	2	JC5581	quanylate cyclase (EC	9.32e+02	945	4	25.0	1860	1	JDMU2	DNA-directed RNA poly	9.32e+02
873	4	25.0	1107	2	C45439	myosin-I, Myr 1b (alt	9.32e+02	946	4	25.0	1867	2	S22775	MOT1 protein - yeast	9.32e+02
874	4	25.0	1110	2	A43253	large tra-1 protein -	9.32e+02	947	4	25.0	1896	1	RNFE2L	DNA-directed RNA poly	9.32e+02
875	4	25.0	1111	2	T01239	hypothetical protein	9.32e+02	948	4	25.0	1918	2	S43719	lactase (EC 3.2.1.108	9.32e+02
876	4	25.0	1113	2	T00271	hypothetical protein	9.32e+02	949	4	25.0	1920	2	S43720	lactase (EC 3.2.1.108	9.32e+02
877	4	25.0	1113	3	JE0315	low-density lipoprote	9.32e+02	950	4	25.0	1926	2	JC4842	DNA-binding nuclear p	9.32e+02
878	4	25.0	1120	2	H71664	transcription-repair	9.32e+02	951	4	25.0	1926	2	S01169	beta-glycosidase comp	9.32e+02
879	4	25.0	1121	2	I38127	phosphoprotein phosph	9.32e+02	952	4	25.0	1927	2	S01168	beta-glycosidase comp	9.32e+02
880	4	25.0	1136	2	A45439	myosin I heavy chain	9.32e+02	953	4	25.0	1928	2	JS0610	beta-galactosidase (E	9.32e+02
881	4	25.0	1138	2	G71554	probable transmembran	9.32e+02	954	4	25.0	1938	2	A37361	probable integral mem	9.32e+02
882	4	25.0	1144	2	A54810	TMV resistance protei	9.32e+02	955	4	25.0	1938	2	JC5421	smooth muscle myosin	9.32e+02
883	4	25.0	1145	2	S13643	PRP22 protein - yeast	9.32e+02	956	4	25.0	1961	1	A61231	myosin heavy chain no	9.32e+02
884	4	25.0	1163	2	G64315	type I restriction en	9.32e+02	957	4	25.0	1972	1	A41604	myosin heavy chain, s	9.32e+02
885	4	25.0	1165	1	GNLJGL	pol polyprotein - gib	9.32e+02	958	4	25.0	1972	2	S68176	TOG protein - human	9.32e+02
886	4	25.0	1166	2	H71609	hypothetical protein	9.32e+02	959	4	25.0	1979	2	C71622	hypothetical protein	9.32e+02
887	4	25.0	1174	2	A39927	DNA-directed RNA poly	9.32e+02	960	4	25.0	1979	2	JW0059	mcpd protein - mouse	9.32e+02
888	4	25.0	1184	2	A55184	fibulin-2 precursor -	9.32e+02	961	4	25.0	1984	2	A44396	P-type cation translo	9.32e+02
889	4	25.0	1184	2	G70600	hypothetical protein	9.32e+02	962	4	25.0	1999	1	S21801	myosin heavy chain, n	9.32e+02
890	4	25.0	1184	2	A39800	calcium-activated pot	9.32e+02	963	4	25.0	2007	1	B43402	myosin heavy chain-B,	9.32e+02
891	4	25.0	1185	2	I58009	gene ERCC5 protein -	9.32e+02	964	4	25.0	2014	2	S46622	probable membrane pro	9.32e+02
892	4	25.0	1186	2	A54439	excision repair prote	9.32e+02	965	4	25.0	2187	2	S60224	polyketide synthase 1	9.32e+02
893	4	25.0	1186	2	S35993	DNA repair protein xp	9.32e+02	966	4	25.0	2231	2	S53416	SENI protein - yeast	9.32e+02
894	4	25.0	1191	2	S27329	DNA topoisomerase (AT	9.32e+02	967	4	25.0	2248	2	D42088	adenylate cyclase (EC	9.32e+02
895	4	25.0	1195	2	A59016	DNA-directed DNA poly	9.32e+02	968	4	25.0	2318	2	A45306	notch 3 protein - mou	9.32e+02
896	4	25.0	1200	2	S68258	DNA polymerase gamma	9.32e+02	969	4	25.0	2325	2	A61208	chondroitin sulfate p	9.32e+02
897	4	25.0	1203	2	S26650	DNA-binding protein 5	9.32e+02	970	4	25.0	2380	2	E71604	hypothetical protein	9.32e+02
898	4	25.0	1211	2	S54500	alpha, alpha-trehalase	9.32e+02	971	4	25.0	2453	2	S60254	nuclear receptor cor	9.32e+02
899	4	25.0	1222	2	S56030	SCP160 protein - yeas	9.32e+02	972	4	25.0	2485	2	H71621	Ser/Thr protein kinas	9.32e+02

973 4 25.0 2549 2 S45340 FKBP-rapamycin-associ 9.32e+02  
974 4 25.0 2549 2 A54837 rapamycin/FKBP12 targ 9.32e+02  
975 4 25.0 2652 1 VF1HB2 genome polyprotein - 9.32e+02  
976 4 25.0 2688 2 I49477 alpha-A-crystallin-bi 9.32e+02  
977 4 25.0 2711 2 A49132 fat facets (faf), spl 9.32e+02  
978 4 25.0 2893 2 A64556 toxin-like outer memb 9.32e+02  
979 4 25.0 2907 2 A57278 fibrillin-2 precursor 9.32e+02  
980 4 25.0 2946 2 T00867 cell division control 9.32e+02  
981 4 25.0 3085 3 T00327 polyprotein - infecti 9.32e+02  
982 4 25.0 3119 2 I49729 HD protein - mouse 9.32e+02  
983 4 25.0 3144 2 A46068 Huntington disease-as 9.32e+02  
984 4 25.0 3187 2 JC5837 364K Golgi complex-as 9.32e+02  
985 4 25.0 3434 1 GNMVMV genome polyprotein - 9.32e+02  
986 4 25.0 3924 2 S37431 ankyrin 2, neuronal 1 9.32e+02  
987 4 25.0 3973 2 B17612 hypothetical protein 9.32e+02  
988 4 25.0 4096 2 A57099 DNA-activated protein 9.32e+02  
989 4 25.0 4128 2 JC6306 protein kinase (EC 2. 9.32e+02  
990 4 25.0 4488 1 R1RHM2 genome polyprotein la 9.32e+02  
991 4 25.0 4544 1 S02392 alpha-2-macroglobulin 9.32e+02  
992 4 25.0 4545 1 S25111 alpha-2-macroglobulin 9.32e+02  
993 4 25.0 4639 2 A54794 dynein heavy chain, c 9.32e+02  
994 4 25.0 4868 2 B54161 ryanodine-binding pro 9.32e+02  
995 4 25.0 5032 2 A35041 ryanodine receptor ty 9.32e+02  
996 4 25.0 5037 2 B35041 ryanodine receptor, s 9.32e+02  
997 4 25.0 6839 2 S57242 twitchin - Caenorhabd 9.32e+02  
998 4 25.0 7962 2 I38346 elastic titin - human 9.32e+02  
999 4 25.0 15281 2 S41309 cyclosporin synthetas 9.32e+02  
1000 4 25.0 26926 1 I38344 titin, cardiac muscle 9.32e+02

ALIGNMENTS

RESULT 1

ENTRY B24479 #type complete  
TITLE calcium-binding protein I - common lancelet  
ALTERNATE\_NAMES sarcoplasmic calcium-binding protein  
ORGANISM #formal\_name Branchiostoma lanceolatum #common\_name common lancelet  
DATE 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 12-Sep-1997

ACCESSIONS B24479  
REFERENCE A90506  
#authors Takagi, T.; Konishi, K.; Cox, J.A.  
#journal Biochemistry (1986) 25:3585-3592  
#title Amino acid sequence of two sarcoplasmic calcium-binding proteins from the protochordate amphioxus.

#accession B24479  
#molecule\_type protein  
#residues 1-183 #label TAK

CLASSIFICATION #superfamily unassigned calmodulin-related proteins; calmodulin repeat homology

KEYWORDS calcium binding; EF hand

FEATURE 134-166 #domain calmodulin repeat homology #label EF4

SUMMARY #length 183 #molecular-weight 21118 #checksum 4156

Query Match 37.5%; Score 6; DB 2; Length 183;  
Best Local Similarity 100.0%; Pred. No. 4.98e-01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69 INKDDV 74  
Qy 2 INKDDV 7  
|||||

RESULT 2

ENTRY A24479 #type complete  
TITLE calcium-binding protein II - common lancelet  
ALTERNATE\_NAMES sarcoplasmic calcium-binding protein  
ORGANISM #formal\_name Branchiostoma lanceolatum #common\_name common lancelet  
DATE 25-Oct-1987 #sequence\_revision 25-Oct-1987 #text\_change 12-Sep-1997

ACCESSIONS A24479  
REFERENCE A90506  
#authors Takagi, T.; Konishi, K.; Cox, J.A.  
#journal Biochemistry (1986) 25:3585-3592  
#title Amino acid sequence of two sarcoplasmic calcium-binding proteins from the protochordate amphioxus.

#accession A24479  
#molecule\_type protein  
#residues 1-183 #label TAK

CLASSIFICATION #superfamily unassigned calmodulin-related proteins; calmodulin repeat homology

KEYWORDS calcium binding; EF hand

FEATURE 134-166 #domain calmodulin repeat homology #label EF4

SUMMARY #length 183 #molecular-weight 21043 #checksum 4182

Query Match 37.5%; Score 6; DB 2; Length 183;  
Best Local Similarity 100.0%; Pred. No. 4.98e-01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 69 INKDDV 74  
Qy 2 INKDDV 7  
|||||

RESULT 3

ENTRY S13183 #type complete  
TITLE calcium-binding protein SCP II - common lancelet  
ORGANISM #formal\_name Branchiostoma lanceolatum #common\_name common lancelet  
DATE 02-Dec-1993 #sequence\_revision 08-Sep-1995 #text\_change 12-Sep-1997

ACCESSIONS S13183  
REFERENCE S13182  
#authors Takagi, T.; Cox, J.A.  
#journal Eur. J. Biochem. (1990) 192:387-399  
#title Amino acid sequences of four isoforms of amphioxus sarcoplasmic calcium-binding proteins.

#cross-references MUID:91006125  
#accession S13183  
#molecule\_type protein  
#residues 1-185 #label TAK

CLASSIFICATION #superfamily unassigned calmodulin-related proteins; calmodulin repeat homology

KEYWORDS alternative splicing; calcium binding; EF hand

FEATURE 6-38 #domain calmodulin repeat homology #label EF1\  
57-89 #domain calmodulin repeat homology #label EF2\  
102-134 #domain calmodulin repeat homology #label EF3\  
136-168 #domain calmodulin repeat homology #label EF4

SUMMARY #length 185 #molecular-weight 21286 #checksum 6749

Query Match 37.5%; Score 6; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 4.98e-01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 71 INKDDV 76  
Qy 2 INKDDV 7  
|||||

RESULT 4

ENTRY S13184 #type complete  
TITLE calcium-binding protein SCP III - common lancelet  
ORGANISM #formal\_name Branchiostoma lanceolatum #common\_name common lancelet  
DATE 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-Aug-1998

ACCESSIONS S13184  
REFERENCE S13182  
#authors Takagi, T.; Cox, J.A.  
#journal Eur. J. Biochem. (1990) 192:387-399  
#title Amino acid sequences of four isoforms of amphioxus

```

#cross-references MUID:91006125
#accession S13184
#status preliminary
#molecule_type protein
#residues 1-185 #label TAK
#note the sequence from Table 4 is inconsistent with that from
Fig. 2 in lacking 43-Lys, 98-ASP, 100-Pro, and 145-Tyr
CLASSIFICATION #superfamily unassigned calmodulin-related proteins;
calmodulin repeat homology
FEATURE
6-38 #domain calmodulin repeat homology #label EF1\
57-89 #domain calmodulin repeat homology #label EF2\
102-134 #domain calmodulin repeat homology #label EF3\
136-168 #domain calmodulin repeat homology #label EF4\
SUMMARY #length 185 #molecular-weight 21386 #checksum 6655
Query Match 37.5%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 INKDDV 76
| | | | |
Qy 2 INKDDV 7

RESULT 5
ENTRY S13182 #type complete
TITLE calcium-binding protein SCP I - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
lancelet
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
07-Aug-1998
ACCESSIONS S13182
REFERENCE Takagi, T.; Cox, J.A.
#authors Eur. J. Biochem. (1990) 192:387-399
#journal Amino acid sequences of four isoforms of amphioxus
#title Sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13182
#status preliminary
#molecule_type protein
#residues 1-185 #label TAK
#note the sequence from Table 4 is inconsistent with that from
Fig. 2 in lacking 43-Lys, 98-ASP, 100-Pro, and 145-Tyr
CLASSIFICATION #superfamily unassigned calmodulin-related proteins;
calmodulin repeat homology
FEATURE
6-38 #domain calmodulin repeat homology #label EF1\
57-89 #domain calmodulin repeat homology #label EF2\
102-134 #domain calmodulin repeat homology #label EF3\
136-168 #domain calmodulin repeat homology #label EF4\
SUMMARY #length 185 #molecular-weight 21418 #checksum 6895
Query Match 37.5%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 INKDDV 76
| | | | |
Qy 2 INKDDV 7

RESULT 6
ENTRY S13185 #type complete
TITLE calcium-binding protein SCP IV - common lancelet
ORGANISM #formal_name Branchiostoma lanceolatum #common_name common
lancelet
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
07-Aug-1998
ACCESSIONS S13185
REFERENCE Takagi, T.; Cox, J.A.
#authors Eur. J. Biochem. (1990) 192:387-399
#journal Amino acid sequences of four isoforms of amphioxus
#title Sarcoplasmic calcium-binding proteins.
#cross-references MUID:91006125
#accession S13185
#status preliminary
#molecule_type protein
#residues 1-185 #label TAK
#note the sequence from Table 4 is inconsistent with that from
Fig. 2 in lacking 43-Lys, 98-ASP, 100-Pro, and 145-Tyr
CLASSIFICATION #superfamily unassigned calmodulin-related proteins;
calmodulin repeat homology
FEATURE
6-38 #domain calmodulin repeat homology #label EF1\
57-89 #domain calmodulin repeat homology #label EF2\
102-134 #domain calmodulin repeat homology #label EF3\
136-168 #domain calmodulin repeat homology #label EF4\
SUMMARY #length 185 #molecular-weight 21417 #checksum 7125
Query Match 37.5%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 71 INKDDV 76
| | | | |
Qy 2 INKDDV 7

RESULT 7
ENTRY B71153 #type complete
TITLE hypothetical protein PH0425 - Pyrococcus horikoshii
ORGANISM #formal_name Pyrococcus horikoshii
DATE 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change
14-Aug-1998
ACCESSIONS B71153
REFERENCE Kwarabavasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.;
Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.;
Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;
Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;
Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.
#journal DNA Res. (1998) 5:55-76
#title Complete sequence and gene organization of the genome of a
hyper-thermophilic archaeobacterium, Pyrococcus horikoshii
OT3.
#cross-references MUID:98344137
#accession B71153
#status preliminary; nucleic acid sequence not shown;
#molecule_type DNA
#residues 1-440 #label KAW
#cross-references GB:AP000002; NID:g3236129; PID:d1030454; PID:g3256828
#experimental_source strain OT3
#note this accession replaces an interim accession for a
sequence replaced by GenBank
GENETICS
#gene PH0425
SUMMARY #length 440 #molecular-weight 50609 #checksum 2190
Query Match 37.5%; Score 6; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 366 INKDDV 371
| | | | |
Qy 2 INKDDV 7

RESULT 8
ENTRY E64414 #type complete
TITLE hypothetical protein MJ0917 - Methanococcus jannaschii
ORGANISM #formal_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
10-Oct-1997

```

```
ACCESSIONS E64414
REFERENCE A64300
#authors Bult, C. J.; White, O.; Olsen, G. J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Ufferback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.P.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal Science (1996) 273:1058-1073
#title Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
#cross-references MUID:96337999
#accession E64414
#status preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-574 #label BUL
#cross-references GB:U67535; GB:L77117; NID:g1591586; PID:g1591590; TIGR:MJ0917; PID:g1510961
GENETICS
#map_position FOR847952-849676
#start_codon GTG
#length 574 #molecular-weight 64141 #checksum 2767
SUMMARY
Query Match 37.5%; Score 6; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 539 EINKDD 544
|||||
Qy 1 EINKDD 6
RESULT 9
ENTRY A40692 #type complete
TITLE Signal recognition particle 72K chain - dog
ORGANISM #formal_name Canis lupus familiaris #common_name dog
DATE 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 10-Sep-1997
ACCESSIONS A40692; S32167
REFERENCE A40692
#authors Lutcke, H.; Prehn, S.; Ashford, A.J.; Remus, M.; Frank, R.; Dobbstein, B.
#journal J. Cell Biol. (1993) 121:977-985
#title Assembly of the 68- and 72-kD proteins of signal recognition particle with 7S RNA.
#cross-references MUID:93273803
#accession A40692
#status preliminary
#molecule_type mRNA; protein
#residues 1-671 #label LUT
#cross-references EMBL:X67813; NID:g297767; PID:g297768
#note sequence extracted from NCBI backbone (NCBIP:132901)
SUMMARY #length 671 #molecular-weight 74493 #checksum 6822
Query Match 37.5%; Score 6; DB 2; Length 671;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 40 INKDDV 45
|||||
Qy 2 INKDDV 7
RESULT 10
ENTRY S42369 #type complete
TITLE Clathrin heavy chain homolog - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 08-Sep-1997
ACCESSIONS S42369
REFERENCE S42368
#authors Smith, A.
#submission submitted to the EMBL Data Library, March 1994
#accession S42369
#status preliminary
#molecule_type DNA
#residues 1-1681 #label SMI
#cross-references EMBL:Z30423; NID:g458479; PID:g458481
GENETICS
#introns 41/3; 82/1; 203/3; 263/3; 550/3; 651/3; 1036/3; 1444/1; 1637/1
CLASSIFICATION #superfamily clathrin heavy chain
SUMMARY #length 1681 #molecular-weight 191541 #checksum 9116
Query Match 37.5%; Score 6; DB 2; Length 1681;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1256 KOVCFS 1261
|||||
Qy 10 KOVCFS 15
RESULT 11
ENTRY S21601 #type complete
TITLE hypothetical protein, 223K - potato virus M
ORGANISM #formal_name potato virus M
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
ACCESSIONS S21601
REFERENCE S21601
#authors Zavriev, S.K.
#submission submitted to the EMBL Data Library, May 1990
#accession S21601
#status preliminary
#molecule_type genomic RNA
#residues 1-1967 #label ZAV
#cross-references EMBL:X53062
#experimental_source Russian wild type
CLASSIFICATION #superfamily eggplant mosaic virus RNA-directed RNA polymerase
SUMMARY #length 1967 #molecular-weight 223207 #checksum 1096
Query Match 37.5%; Score 6; DB 2; Length 1967;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1324 EINKDD 1329
|||||
Qy 1 EINKDD 6
RESULT 12
ENTRY PN0093 #type complete
TITLE genome polyprotein - potato virus M (strain Russian)
ALTERNATE_NAMES RNA nucleotidyltransferase (RNA-directed); RNA replicase
CONTAINS RNA-directed RNA polymerase (EC 2.7.7.48)
ORGANISM #formal_name potato virus M
DATE 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 29-May-1998
ACCESSIONS PN0093; A54333
REFERENCE PN0093
#authors Zavriev, S.K.; Kanyka, K.V.; Levay, K.E.
#journal Mol. Biol. (Mosk.) (1991) 25:761-769
#title The complete nucleotide sequence of potato virus M genomic RNA.
#cross-references MUID:92049299
#accession PN0093
#molecule_type genomic RNA
#residues 1-1968 #label ZAV
#cross-references GB:X53062; NID:g61291; PID:g61292
```



```

##note      the authors translated the codons GAA for residues 79
            and 288 as Ile, GTG for residue 529 as Phe, CAC for
            residues 548 and 630 as Asn, AGA for residue 646 as
            Pro, CCA for residue 1251 as Val, and ATG for residues
            799, 1261, and 1366 as Asn

REFERENCE    A54333
#authors     Zavriev, S.K.; Kanyuka, K.V.; Levay, K.E.
#journal     J. Gen. Virol. (1991) 72:9-14
#title       The genome organization of potato virus M RNA.
#accession   A54333
#molecule_type genomic RNA
##residues   1-1968 ##label ZA2
##cross-references EMBL:X53062; NID:g61291; PID:g61292
##note      authors translated the codon GAA for residue 79 as Ile,
            GAA for residue 288 as Ile, CCG for residue 356 as
            Arg, GTG for residue 529 as Phe, CAC for residue 548
            as Asn, CAC for residue 630 as Asn, CCC for residue
            645 as Arg, AGA for residue 646 as Pro, and ATG for
            residue 779 as Asn
##note      authors translated the codon CGA for residue 882 as Pro,
            CCA for residue 1251 as Val, ATG for residue 1261 as
            Asn, ATG for residue 1366 as Asn, TTT for residue 1380
            as Glu, CAC for residue 1532 as Asn, CCC for residue
            1578 as Arg, CGG for residue 1635 as Pro, and CCC for
            residue 1710 as Arg

CLASSIFICATION #superfamily eggplant mosaic virus RNA-directed RNA
                polymerase
KEYWORDS       ATP; nucleotidyltransferase; P-loop; RNA biosynthesis; RNA
                replication
FEATURE        1166-1173   #region nucleotide-binding motif A (P-loop)\
                1241-1246   #region nucleotide-binding motif B\
                1172        #binding_site ATP (Lys) #status predicted
SUMMARY        #length 1968 #molecular-weight 223384 #checksum 4558

Query Match    37.5%; Score 6; DB 1; Length 1968;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1325 EINKDD 1330
|||||
QY 1 EINKDD 6

RESULT 13
ENTRY   A24420      #type complete
TITLE   notch protein - fruit fly (Drosophila melanogaster)
ALTERNATE_NAMES #neurogenic repetitive locus protein
ORGANISM #formal_name Drosophila melanogaster
DATE     30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change
        07-Aug-1998
ACCESSIONS A24420; A24768; S09358; A05267
REFERENCE   A24420
#authors    Kidd, S.; Kelley, M.R.; Young, M.W.
#journal     Mol. Cell. Biol. (1986) 6:3094-3108
#cross-references MUID:87064624
#accession   A24420
#molecule_type DNA
##residues   1-2703 ##label K1D
##cross-references GB:R03508; NID:g157991; PID:g157993
REFERENCE   A24768
#authors     Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
#journal     Cell (1985) 43:567-581
#cross-references MUID:86079539
#accession   A24768
#molecule_type mRNA
##residues   1-48, '1', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A',
            268-872, 'R', 874-958, 'R', 960-1970, 'PH', 1973-2256, 'G',
            2258-2264, 'V', 2266-2406, 'R', 2408-2444, 'L', 2446-2703
            ##label WHA1
##note      the authors translated the codon ATC for residue 49 as
            Thr, ATT for residue 2044 as Arg, GTA for residue 2265
            as Ala, CGC for residue 2407 as His, and CTT for

```

```

            residue 2445 as Arg
REFERENCE    S09358
#authors     Tautz, D.
#journal     Nucleic Acids Res. (1989) 17:6463-6471
#title       Hypervariability of simple sequences as a general source for
            polymorphic DNA markers.
#cross-references MUID:89385974
#accession   S09358
#molecule_type DNA
##residues   2505-2551, 'QQQQ', 2552-2576, 'E', 2578-2604 ##label TAU
REFERENCE    A05267
#authors     Wharton, K.A.; Vedvobnick, B.; Finnerty, V.G.;
            Artavanis-Tsakonas, S.
#journal     Cell (1985) 40:55-62
#title       opa: a novel family of transcribed repeats shared by the
            Notch locus and other developmentally regulated loci in D.
            melanogaster.
#cross-references MUID:85099329
#accession   A05267
#molecule_type DNA
##residues   2504-2576, 'E', 2578-2611 ##label WHA2
GENETICS
#gene         notch; opa
#map_position 8.96-9.36
#introns      53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3
CLASSIFICATION #superfamily notch protein; ankyrin repeat homology; EGF
                homology
KEYWORDS       differentiation; tandem repeat; transmembrane protein
FEATURE        27-43      #domain transmembrane #status predicted #label TMM1\
                568-599   #domain EGF homology #label EGF\
                1746-1762 #domain transmembrane #status predicted #label TMM2\
                1950-1982 #domain ankyrin repeat homology #label AN1\
                1983-2015 #domain ankyrin repeat homology #label AN2\
                1988-2004 #domain transmembrane #status predicted #label TMM3\
                2037-2049 #domain ankyrin repeat homology #label AN3\
                2050-2082 #domain ankyrin repeat homology #label AN4\
                2083-2115 #domain ankyrin repeat homology #label AN5\
                2538-2568 #region glutamine-rich\
                2538-2568 #domain neurogenic repetitive element #status predicted
                #label OPA
SUMMARY        #length 2703 #molecular-weight 288876 #checksum 6404

Query Match    37.5%; Score 6; DB 2; Length 2703;
Best Local Similarity 100.0%; Pred. No. 4.98e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1257 EINKDD 1262
|||||
QY 1 EINKDD 6

RESULT 14
ENTRY   A45681      #type complete
TITLE   hypothetical 7K protein (gene 61.5 protein) - phage T4
ALTERNATE_NAMES #formal_name phage T4
ORGANISM #host Echerichia coli
#note     21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change
        29-Jan-1999
ACCESSIONS A45681; JS0558
REFERENCE   A45681
#authors     Selick, H.E.; Stormo, G.D.; Dyson, R.L.; Alberts, B.M.
#journal     J. Virol. (1993) 67:2305-2316
#title       Analysis of five presumptive protein-coding sequences
            clustered between the primosome genes, 41 and 61, of
            bacteriophages T4, T2, and T6.
#cross-references MUID:93188183
#accession   A45681
#status      Preliminary
#molecule_type DNA
##residues   1-60 ##label SEL

```

##cross-references GB:S57514; NID:g298518; PID:g298519  
sequence extracted from NCBI backbone (NCBIN:128289,  
NCBIP:128290)

## GENETICS

#map\_position 20.378-20.558  
SUMMARY #length 60 #molecular-weight 7027 #checksum 5034

Query Match 31.3%; Score 5; DB 2; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.76e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 KDDVQ 30  
|||||  
Qy 4 KDDVQ 8

## RESULT 15

ENTRY E64356 #type complete  
TITLE conserved hypothetical protein MJ0453 - Methanococcus  
jannaschii  
ORGANISM #formal\_name Methanococcus jannaschii  
DATE 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change  
18-Sep-1998  
ACCESSIONS E64356  
REFERENCE A64300  
#authors Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,  
R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,  
R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;  
Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;  
Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,  
J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,  
J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,  
K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk, J.C.  
H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
#journal Science (1996) 273:1058-1073  
#title Complete genome sequence of the methanogenic archaeon,  
Methanococcus jannaschii.  
#cross-references MUID:96337999  
#accession E64356  
#status preliminary; nucleic acid sequence not shown;  
translation not shown  
##molecule\_type DNA  
##residues 1-107 #label BUL  
##cross-references GB:U67496; GB:L77117; NID:g1591148; PID:g1591157;  
TIGR:MJ0453; PID:g1510528

## GENETICS

#map\_position REV405287-404964  
CLASSIFICATION #superfamily Methanococcus jannaschii conserved hypothetical  
protein MJ0453  
SUMMARY #length 107 #molecular-weight 12603 #checksum 8821

Query Match 31.3%; Score 5; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 2.76e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 EINKD 85  
|||||  
Qy 1 EINKD 5

Search completed: Sat Aug 28 15:13:13 1999  
Job time : 47 secs.

\*\*\*\*\*  
WISOR  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 15:10:13 1999; Maspar time 3.21 Seconds  
Tabular output not generated.  
141.093 Million cell updates/sec

Title: >US-09-049-696-49  
Description: (1-16) from US09049696.pgp  
Perfect Score: 16  
Sequence: 1 EINKDDVQHKQVCFSR 16

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 2.366; Variance 0.343; scale 6.891

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	37.5	185	1 SCPI_BRALA	SARCOPLASMIC CALCIUM-B	2.43e-01
2	6	37.5	185	1 SCPI_BRALA	SARCOPLASMIC CALCIUM-B	2.43e-01
3	6	37.5	670	1 SR72_CANFA	SIGNAL RECOGNITION PAR	2.43e-01
4	6	37.5	670	1 SR72_HUMAN	SIGNAL RECOGNITION PAR	2.43e-01
5	6	37.5	877	1 MGR3_HUMAN	METABOTROPIC GLUTAMATE	2.43e-01
6	6	37.5	1681	1 CLH_CAEEL	PROBABLE CLATHRIN HEAV	2.43e-01
7	6	37.5	1967	1 RPOC_PVMR	RNA REPLICATION PROTEI	2.43e-01
8	6	37.5	2703	1 NOVC_DROME	NEUROGENIC LOCUS NOTCH	2.43e-01
9	5	31.3	60	1 Y02B_BPT4	HYPOTHETICAL 7.0 KD PR	1.64e+01
10	5	31.3	107	1 Y453_METJA	HYPOTHETICAL PROTEIN M	1.64e+01
11	5	31.3	133	1 YJPI_YEAST	HYPOTHETICAL 15.2 KD P	1.64e+01
12	5	31.3	183	1 Y391_METJA	HYPOTHETICAL PROTEIN M	1.64e+01
13	5	31.3	219	1 TPIS_METJA	PUTATIVE TRIOSPHOSPHA	1.64e+01
14	5	31.3	227	1 PINT_BRARE	PROTEIN-L-ISOASPARTATE	1.64e+01
15	5	31.3	228	1 SEM5_CAEEL	SEX MUSCLE ABNORMAL PR	1.64e+01
16	5	31.3	238	1 NEUM_HUMAN	NEUROMODULIN (AXONAL M	1.64e+01
17	5	31.3	239	1 YGIP_YEAST	HYPOTHETICAL 27.6 KD P	1.64e+01
18	5	31.3	243	1 YC22_METJA	HYPOTHETICAL PROTEIN M	1.64e+01
19	5	31.3	265	1 APAL_TUGB	APOLIPOPROTEIN A-I PRE	1.64e+01
20	5	31.3	267	1 YCHA_HAEIN	HYPOTHETICAL PROTEIN H	1.64e+01
21	5	31.3	279	1 Y43D_MYCPN	HYPOTHETICAL LIPOPROTE	1.64e+01
22	5	31.3	281	1 BLAC_STAAP	BETA-LACTAMASE PRECURS	1.64e+01
23	5	31.3	289	1 YADA_YEAST	HYPOTHETICAL 33.2 KD P	1.64e+01

AUTOPHAGY PROTEIN AP05 1.64e+01  
RPIR PROTEIN. 1.64e+01  
HYPOTHETICAL TRANSCRIP 1.64e+01  
ALKANAL MONOOXYGENASE 1.64e+01  
DNA-DIRECTED RNA POLYM 1.64e+01  
LDO/AO TRANSPORT SYSTE 1.64e+01  
CHLOROPHYLLIDE REDUCTA 1.64e+01  
CHLOROPHYLLIDE REDUCTA 1.64e+01  
LICA PROTEIN. 1.64e+01  
SODIUM-DEPENDENT DOPAM 1.64e+01  
VANCOMYCIN/TEICOPLAMIN 1.64e+01  
ALL PROTEIN. 1.64e+01  
PECTIN LYASE D PRECURS 1.64e+01  
HYPOTHETICAL 44.6 KD P 1.64e+01  
PROBABLE NADH-DEPENDEN 1.64e+01  
PROTEIN PHOSPHATASE 2C 1.64e+01  
PROTEIN PHOSPHATASE 2C 1.64e+01  
PROTEIN PHOSPHATASE 2C 1.64e+01  
NIFS-LIKE PROTEIN. 1.64e+01  
SEPTIN 2. 1.64e+01  
PREPROTEIN TRANSLOCASE 1.64e+01  
HEAD VERTEX PROTEIN GP 1.64e+01  
DNA REPAIR PROTEIN RAD 1.64e+01  
ASPARTATE AMINOTRANSFE 1.64e+01  
HYPOTHETICAL PROTEIN M 1.64e+01  
VASOACTIVE INTESTINAL 1.64e+01  
VASOACTIVE INTESTINAL 1.64e+01  
HYPOTHETICAL 54.4 KD P 1.64e+01  
55 KD ERYTHROCYTE MEMB 1.64e+01  
55 KD ERYTHROCYTE MEMB 1.64e+01  
FASCIN. 1.64e+01  
ZINC FINGER PROTEIN HM 1.64e+01  
TRANSCRIPTION FACTOR B 1.64e+01  
HYPOTHETICAL 58.0 KD P 1.64e+01  
ATP SYNTHASE ALPHA CHA 1.64e+01  
DIHYDROLIPOAMIDE ACETY 1.64e+01  
TRYPTOPHAN 2-MONOOXYGE 1.64e+01  
RAB GERANYLGERANYLTRAN 1.64e+01  
THIAMINE METABOLISM RE 1.64e+01  
ELONGATION FACTOR 1 AL 1.64e+01  
SODIUM-DEPENDENT DOPAM 1.64e+01  
SODIUM-DEPENDENT DOPAM 1.64e+01  
MANNOSYL-OLIGOSACCHARI 1.64e+01  
MANNOSYL-OLIGOSACCHARI 1.64e+01  
TRANSLATION INITIATION 1.64e+01  
ALPHA-AMYLASE 1 (EC 3. 1.64e+01  
ZINC FINGER PROTEIN MS 1.64e+01  
HYPOTHETICAL 78.8 KD P 1.64e+01  
HOMOACONITASE PRECURSO 1.64e+01  
DNA GYRASE SUBUNIT A ( 1.64e+01  
METABOTROPIC GLUTAMATE 1.64e+01  
COATOMER GAMMA SUBUNIT 1.64e+01  
PYRUVATE, PHOSPHATE DIK 1.64e+01  
PYRUVATE, PHOSPHATE DIK 1.64e+01  
HYPOTHETICAL 108.7 KD 1.64e+01  
PROTOTOVILLIN (100 KD AC 1.64e+01  
HYPOTHETICAL 127.3 KD 1.64e+01  
HYPOTHETICAL 128.8 KD 1.64e+01  
HYPOTHETICAL 136.3 KD 1.64e+01  
DNA-DIRECTED RNA POLYM 1.64e+01  
DNA-DIRECTED RNA POLYM 1.64e+01  
BETA-N-ACETYLHEXOSAMIN 1.64e+01  
DNA-DIRECTED RNA POLYM 1.64e+01  
XANTHINE DEHYDROGENASE 1.64e+01  
VITELLOGENIN 6 PRECURS 1.64e+01  
INTEGRIN ALPHA CHAIN-L 1.64e+01  
ACETYL-COA CARBOXYLASE 1.64e+01  
ADENOMATOUS POLYPOSIIS 1.64e+01  
UNKNOWN PROTEIN FROM 2 6.35e+02  
NAD-REDUCING HYDROGENA 6.35e+02  
NONSTRUCTURAL 4.8 KD P 6.35e+02

97	4	25.0	53	1	AIAL_HORSE	ALPHA-1-ANTITRYPTEINASE	6.35e+02	170	4	25.0	149	1	RL9_MYCPN	50S RIBOSOMAL PROTEIN	6.35e+02
98	4	25.0	56	1	Y821_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	171	4	25.0	149	1	MLE_BRAFL	MYOSIN, ESSENTIAL LIGH	6.35e+02
99	4	25.0	58	1	TAT_HV1B5	TAT PROTEIN (TRANSACTI	6.35e+02	172	4	25.0	149	1	WLE3_CHICK	MYOSIN LIGHT CHAIN 3,	6.35e+02
100	4	25.0	61	1	3BH5_VARV	TRUNCATED 3-BETA HYDRO	6.35e+02	173	4	25.0	150	1	WLEF_BACSU	PUTATIVE LOW MOLECULAR	6.35e+02
101	4	25.0	71	1	IFIL_BACSU	TRANSLATION INITIATION	6.35e+02	174	4	25.0	151	1	YOM1_PHOPR	PUTATIVE ATP-DEPENDENT	6.35e+02
102	4	25.0	75	1	YXAI_PORPU	HYPOTHETICAL 8.3 KD PR	6.35e+02	175	4	25.0	152	1	VNS2_PVM	NONSTRUCTURAL PROTEIN	6.35e+02
103	4	25.0	76	1	YC34_PORPU	HYPOTHETICAL 8.9 KD PR	6.35e+02	176	4	25.0	153	1	YN21_YEAST	HYPOTHETICAL 17.2 KD P	6.35e+02
104	4	25.0	81	1	Y78A_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	177	4	25.0	154	1	GLB2_NIBPR	MYOGLOBIN (GLOBIN, BO	6.35e+02
105	4	25.0	83	1	YXBA_BACSU	HYPOTHETICAL 9.8 KD PR	6.35e+02	178	4	25.0	156	1	MLR_CHLNI	MYOSIN REGULATORY LIGH	6.35e+02
106	4	25.0	86	1	TAT_HV1BR	TAT PROTEIN (TRANSACTI	6.35e+02	179	4	25.0	156	1	SMPB_BACSU	SMALL PROTEIN B HOMOLO	6.35e+02
107	4	25.0	86	1	TAT_HV112	TAT PROTEIN (TRANSACTI	6.35e+02	180	4	25.0	156	1	IF52_YEAST	INITIATION FACTOR 5A-2	6.35e+02
108	4	25.0	86	1	TAT_HVIND	TAT PROTEIN (TRANSACTI	6.35e+02	181	4	25.0	156	1	IF51_YEAST	INITIATION FACTOR 5A-1	6.35e+02
109	4	25.0	86	1	TAT_HV1H2	TAT PROTEIN (TRANSACTI	6.35e+02	182	4	25.0	160	1	YEMO_YEAST	HYPOTHETICAL 18.3 KD P	6.35e+02
110	4	25.0	90	1	YGX_ECOLI	10.8 KD PROTEIN IN MUT	6.35e+02	183	4	25.0	162	1	DYR_LACCA	DIHYDROFOLATE REDUCTAS	6.35e+02
111	4	25.0	91	1	YJFY_ECOLI	HYPOTHETICAL 10.1 KD P	6.35e+02	184	4	25.0	165	1	HSCB_BUCAP	CHAPERONE PROTEIN HSCB	6.35e+02
112	4	25.0	92	1	EXSE_TRICO	11 KD EXCRETORY-SECRET	6.35e+02	185	4	25.0	167	1	YXG3_YEAST	HYPOTHETICAL 19.0 KD P	6.35e+02
113	4	25.0	94	1	GLHA_THOUB	GLYCOPROTEIN HORMONES	6.35e+02	186	4	25.0	167	1	THIX_HAEIN	THIOREDOXIN-LIKE PROTE	6.35e+02
114	4	25.0	95	1	YG05_MOUSE	HYPOTHETICAL PROTEIN L	6.35e+02	187	4	25.0	169	1	YKH0_YEAST	HYPOTHETICAL 19.8 KD P	6.35e+02
115	4	25.0	96	1	GLHA_RABIT	GLYCOPROTEIN HORMONES	6.35e+02	188	4	25.0	169	1	Y769_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
116	4	25.0	96	1	GLHA_STRCA	GLYCOPROTEIN HORMONES	6.35e+02	189	4	25.0	169	1	MOAB_ECOLI	MOLYBDENUM COFACTOR BI	6.35e+02
117	4	25.0	98	1	CVTA_HUMAN	CYSTATIN A (STEFTN A)	6.35e+02	190	4	25.0	172	1	YHBN_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02
118	4	25.0	101	1	TAT_HV1C4	TAT PROTEIN (TRANSACTI	6.35e+02	191	4	25.0	173	1	GLBC_NIBPR	GLOBIN, CUTICULAR ISOF	6.35e+02
119	4	25.0	101	1	TAT_HV1JR	TAT PROTEIN (TRANSACTI	6.35e+02	192	4	25.0	174	1	KAD_NEIPO	ADENYLATE KINASE (EC 2	6.35e+02
120	4	25.0	101	1	TAT_HV1S3	TAT PROTEIN (TRANSACTI	6.35e+02	193	4	25.0	174	1	VGAM_BPMU	HOST-NUCLEASE INHIBITO	6.35e+02
121	4	25.0	103	1	YD29_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	194	4	25.0	174	1	KAD_NEILA	ADENYLATE KINASE (EC 2	6.35e+02
122	4	25.0	105	1	Y903_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	195	4	25.0	174	1	KAD_NEICI	ADENYLATE KINASE (EC 2	6.35e+02
123	4	25.0	108	1	RODL_NEUCR	HYDROPHOBIN PRECURSOR	6.35e+02	196	4	25.0	175	1	YXG2_YEAST	HYPOTHETICAL 19.2 KD P	6.35e+02
124	4	25.0	108	1	GLH1_ONCKE	GLYCOPROTEIN HORMONES	6.35e+02	197	4	25.0	176	1	APT_BORBU	ADENINE PHOSPHORIBOSYL	6.35e+02
125	4	25.0	109	1	Y353_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	198	4	25.0	178	1	FLIL_BORBU	FLAGELLAR FLIL PROTEIN	6.35e+02
126	4	25.0	109	1	Y353_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02	199	4	25.0	179	1	E320_ADE03	EARLY E3 20.1 KD GLYCO	6.35e+02
127	4	25.0	109	1	Y639_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	200	4	25.0	179	1	170K_TRVPS	POTENTIAL 170 KD PROTE	6.35e+02
128	4	25.0	112	1	WHEP_CAMDR	WHEY PROTEIN.	6.35e+02	201	4	25.0	180	1	LYTB_PSEFL	LYTB PROTEIN HOMOLOG (	6.35e+02
129	4	25.0	112	1	ZG16_XENLA	GASTRULA ZINC FINGER P	6.35e+02	202	4	25.0	181	1	HSLV_BACSU	HEAT SHOCK PROTEIN HSL	6.35e+02
130	4	25.0	113	1	YXG1_YEAST	HYPOTHETICAL 13.5 KD P	6.35e+02	203	4	25.0	182	1	YHBA_ECOLI	HYPOTHETICAL 20.0 KD P	6.35e+02
131	4	25.0	114	1	V840_BPT4	HEAD FORMATION PROTEIN	6.35e+02	204	4	25.0	185	1	DPL_MOUSE	POLYPOSIS LOCUS PROTEI	6.35e+02
132	4	25.0	117	1	GLHA_MORSA	GLYCOPROTEIN HORMONES	6.35e+02	205	4	25.0	187	1	YDAL_ECOLI	HYPOTHETICAL 21.5 KD P	6.35e+02
133	4	25.0	117	1	HIA2_DICDI	DNA-DIRECTED RNA POLYM	6.35e+02	206	4	25.0	187	1	YQGN_BACSU	HYPOTHETICAL 21.4 KD P	6.35e+02
134	4	25.0	117	1	RPBY_HUMAN	HIS-2 DIRECTED RNA POLYM	6.35e+02	207	4	25.0	188	1	KLK3_RAT	GLANDULAR KALLIKREIN 3	6.35e+02
135	4	25.0	117	1	GLHA_ACALA	GLYCOPROTEIN HORMONES	6.35e+02	208	4	25.0	190	1	VIS1_RAT	VISININ-LIKE PROTEIN 1	6.35e+02
136	4	25.0	118	1	GLHA_CTEID	GLYCOPROTEIN HORMONES	6.35e+02	209	4	25.0	190	1	VIS1_HUMAN	VISININ-LIKE PROTEIN 1	6.35e+02
137	4	25.0	118	1	GLHA_HPYMO	GLYCOPROTEIN HORMONES	6.35e+02	210	4	25.0	191	1	G25P_HUMAN	G25K GTP-BINDING PROTE	6.35e+02
138	4	25.0	119	1	PLI1_SALT1	FIMBRIAL PROTEIN PRECU	6.35e+02	211	4	25.0	192	1	UBC1_CAEEL	UBIQUITIN-CONJUGATING	6.35e+02
139	4	25.0	120	1	GLHA_MOUSE	GLYCOPROTEIN HORMONES	6.35e+02	212	4	25.0	194	1	YGG8_YEAST	PUTATIVE MITOCHONDRIAL	6.35e+02
140	4	25.0	120	1	GLHA_MELGA	GLYCOPROTEIN HORMONES	6.35e+02	213	4	25.0	195	1	YXSC_BACSU	HYPOTHETICAL GTP-BINDI	6.35e+02
141	4	25.0	120	1	GLHA_HORSE	GLYCOPROTEIN HORMONES	6.35e+02	214	4	25.0	196	1	ZG64_XENLA	GASTRULA ZINC FINGER P	6.35e+02
142	4	25.0	120	1	GLHA_BOVIN	GLYCOPROTEIN HORMONES	6.35e+02	215	4	25.0	196	1	ZG3_XENLA	GASTRULA ZINC FINGER P	6.35e+02
143	4	25.0	120	1	GLHA_PIG	GLYCOPROTEIN HORMONES	6.35e+02	216	4	25.0	197	1	Y916_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02
144	4	25.0	120	1	GLH2_RAT	GLYCOPROTEIN HORMONES	6.35e+02	217	4	25.0	198	1	HIS5_METJA	AMIDOTRANSFERASE HISH	6.35e+02
145	4	25.0	120	1	GLHA_SHEEP	GLYCOPROTEIN HORMONES	6.35e+02	218	4	25.0	199	1	Y599_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
146	4	25.0	122	1	YBBC_ECOLI	HYPOTHETICAL 14.2 KD P	6.35e+02	219	4	25.0	200	1	YAR1_YEAST	ANKYRIN REPEAT-CONTAIN	6.35e+02
147	4	25.0	123	1	Y055_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02	220	4	25.0	200	1	REP2_ZYGE	TRANS-ACTING FACTOR C	6.35e+02
148	4	25.0	124	1	Y878_BPP21	HYPOTHETICAL 13.5 KD P	6.35e+02	221	4	25.0	201	1	YMW3_YEAST	HYPOTHETICAL 23.2 KD P	6.35e+02
149	4	25.0	127	1	WAP_RABIT	WHEY ACIDIC PROTEIN PR	6.35e+02	222	4	25.0	201	1	COX3_SYNVU	CYTOCHROME C OXIDASE P	6.35e+02
150	4	25.0	130	1	YHIT_CAEEL	HYPOTHETICAL HIT-LIKE	6.35e+02	223	4	25.0	201	1	RANG_YEAST	RAN-SPECIFIC GTPASE-AC	6.35e+02
151	4	25.0	130	1	RL3_PIG	60S RIBOSOMAL PROTEIN	6.35e+02	224	4	25.0	203	1	IF6_SULAC	PUTATIVE TRANSLATION I	6.35e+02
152	4	25.0	133	1	YHNG_ECOLI	HYPOTHETICAL 15.1 KD P	6.35e+02	225	4	25.0	204	1	LAF4_MOUSE	LAF-4 PROTEIN (LYMPHOI	6.35e+02
153	4	25.0	138	1	ATPE_BUCAP	ATP SYNTHASE EPSILON C	6.35e+02	226	4	25.0	206	1	Y140_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
154	4	25.0	138	1	PA2B_TRIFL	PHOSPHOLIPASE A2, BASI	6.35e+02	227	4	25.0	208	1	GRPE_STAAU	GRPE PROTEIN (HSP20)	6.35e+02
155	4	25.0	140	1	Y337_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	228	4	25.0	209	1	GLOB_TOBAC	FLORAL HOMEOTIC PROTEI	6.35e+02
156	4	25.0	141	1	Y496_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	229	4	25.0	210	1	Y753_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
157	4	25.0	141	1	PHCO_BPT4	17 KD PROHEAD CORE PRO	6.35e+02	230	4	25.0	211	1	YNO8_CAEEL	HYPOTHETICAL 23.7 KD P	6.35e+02
158	4	25.0	141	1	RL11_HELPY	50S RIBOSOMAL PROTEIN	6.35e+02	231	4	25.0	211	1	PMGY_SCHPO	PHOSPHOGLYCERATE MUTAS	6.35e+02
159	4	25.0	142	1	CYTD_HUMAN	CYSTATIN D PRECURSOR.	6.35e+02	232	4	25.0	211	1	GSPI_XANCP	GENERAL SECRETION PATH	6.35e+02
160	4	25.0	144	1	Y331_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	233	4	25.0	213	1	HIS1_BACSU	ATP PHOSPHORIBOSYLFRAN	6.35e+02
161	4	25.0	144	1	YYS_YEAST	HYPOTHETICAL PROTEIN I	6.35e+02	234	4	25.0	213	1	DEF7_DERFA	MITE ALLERGEN DER F 7	6.35e+02
162	4	25.0	145	1	Y613_SYNV3	HYPOTHETICAL 15.8 KD P	6.35e+02	235	4	25.0	214	1	YDL8_SCHPO	HYPOTHETICAL 23.9 KD P	6.35e+02
163	4	25.0	146	1	YD02_VACCC	PROTEIN D2.	6.35e+02	236	4	25.0	215	1	TRPF_BACSU	N-(5'-PHOSPHORIBOSYL)A	6.35e+02
164	4	25.0	146	1	YAAR_BACSU	ALLOGRAFT INFLAMMATORY	6.35e+02	237	4	25.0	215	1	KAD_NEIGO	ADENYLATE KINASE (EC 2	6.35e+02
165	4	25.0	146	1	ATFL_PIG	HYPOTHETICAL 16.7 KD P	6.35e+02	238	4	25.0	215	1	RB14_RAT	RAS-RELATED PROTEIN RA	6.35e+02
166	4	25.0	147	1	RL13_MYCLE	50S RIBOSOMAL PROTEIN	6.35e+02	239	4	25.0	215	1	DER7_DERPT	MITE ALLERGEN DER P 7	6.35e+02
167	4	25.0	148	1	YHT6_YEAST	HYPOTHETICAL 17.0 KD P	6.35e+02	240	4	25.0	215	1	EF1B_WHEAT	ELONGATION FACTOR 1 BE	6.35e+02
168	4	25.0	148	1	Y452_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	241	4	25.0	217	1	RIC2_ORYSA	RAS-RELATED PROTEIN RI	6.35e+02
169	4	25.0	148	1	RL19_METH	50S RIBOSOMAL PROTEIN	6.35e+02	242	4	25.0	217	1	FK21_NEUCR	FK506-BINDING PROTEIN	6.35e+02

243	4	25.0	1	GRB2_XENLA	GROWTH FACTOR RECEPTOR	6.35e+02	316	4	25.0	1	YAFB_ECOLI	HYPOTHETICAL OXIDOREDU	6.35e+02	
244	4	25.0	1	GRB2_HUMAN	GROWTH FACTOR RECEPTOR	6.35e+02	317	4	25.0	1	CLCR_RAT	CALDECRRIN PRECURSOR (E	6.35e+02	
245	4	25.0	1	Y381_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02	318	4	25.0	1	YDR8_SCHPO	HYPOTHETICAL 30.5 KD P	6.35e+02	
246	4	25.0	1	YB9J_YEAST	HYPOTHETICAL RAS-RELAT	6.35e+02	319	4	25.0	270	1	YDLJ_SCHPO	HYPOTHETICAL 31.5 KD P	6.35e+02
247	4	25.0	1	REHY_HORVU	PEROXIREDOXIN (REHYDRI	6.35e+02	320	4	25.0	271	1	YDJ0_ECOLI	HYPOTHETICAL 30.7 KD P	6.35e+02
248	4	25.0	1	REHY_YEAST	HYPOTHETICAL 25.9 KD P	6.35e+02	321	4	25.0	271	1	YDJB_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02
249	4	25.0	1	YMW7_YEAST	HYPOTHETICAL 25.6 KD P	6.35e+02	322	4	25.0	272	1	YDV3_CAEEL	HYPOTHETICAL 31.9 KD P	6.35e+02
250	4	25.0	1	CTEP_CLOAB	BUTYRATE-ACETOACETATE	6.35e+02	323	4	25.0	272	1	NIF2_CLOPA	NITROGENASE IRON PROTE	6.35e+02
251	4	25.0	1	FLIP_BACSU	FLAGELLAR BIOSYNTHETIC	6.35e+02	324	4	25.0	274	1	Y179_MYCGE	HYPOTHETICAL ABC TRANS	6.35e+02
252	4	25.0	1	SV5	NONSTRUCTURAL PROTEIN	6.35e+02	325	4	25.0	276	1	PROC_ARATH	PYRROLINE-5-CARBOXYLAT	6.35e+02
253	4	25.0	1	V101_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	326	4	25.0	277	1	YKC3_YEAST	HYPOTHETICAL 32.3 KD P	6.35e+02
254	4	25.0	1	YKGH_ECOLI	HYPOTHETICAL 25.6 KD P	6.35e+02	327	4	25.0	278	1	PROC_ACTCH	PYRROLINE-5-CARBOXYLAT	6.35e+02
255	4	25.0	1	YKGF_ECOLI	HYPOTHETICAL 25.4 KD P	6.35e+02	328	4	25.0	279	1	VRXA_LAMBD	REXA PROTEIN.	6.35e+02
256	4	25.0	1	KADL_YEAST	ADENYLATE KINASE CYTOS	6.35e+02	329	4	25.0	280	1	ZG46_ZENLA	GASTRULA ZINC FINGER P	6.35e+02
257	4	25.0	1	ZO61_XENLA	OOCYTE ZINC FINGER PRO	6.35e+02	330	4	25.0	280	1	CHIA_MAIZE	ENDOCHITININASE A PRECUR	6.35e+02
258	4	25.0	1	YF59_SVNY3	HYPOTHETICAL 24.9 KD P	6.35e+02	331	4	25.0	283	1	Y43C_MYCPN	HYPOTHETICAL LIPOPROTE	6.35e+02
259	4	25.0	1	Y703_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	332	4	25.0	284	1	Y309_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
260	4	25.0	1	Y069_HUMAN	HYPOTHETICAL PROTEIN K	6.35e+02	333	4	25.0	284	1	COX3_LEITA	CYTOCHROME C OXIDASE P	6.35e+02
261	4	25.0	1	NRFB_HAEIN	CYTOCHROME C-TYPE PROT	6.35e+02	334	4	25.0	284	1	TPM2_DROME	TROPOMYOSIN 2, MUSCLE	6.35e+02
262	4	25.0	1	RAN_GIALA	GTP-BINDING NUCLEAR PR	6.35e+02	335	4	25.0	285	1	Y007_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02
263	4	25.0	1	Y132_NPVOP	HYPOTHETICAL 26.0 KD P	6.35e+02	336	4	25.0	285	1	YDH5_SCHPO	HYPOTHETICAL 32.8 KD P	6.35e+02
264	4	25.0	1	VHEL_LSV	PROBABLE HELICASE (ORF	6.35e+02	337	4	25.0	285	1	YHMF_METPE	HYPOTHETICAL 32.2 KD P	6.35e+02
265	4	25.0	1	IF6_METJA	PUTATIVE TRANSLATION I	6.35e+02	338	4	25.0	285	1	MENB_HAEIN	NAPHTHOATE SYNTHASE (E	6.35e+02
266	4	25.0	1	TRMD_HELPY	TRNA (GUANINE-NI)-METH	6.35e+02	339	4	25.0	285	1	FLA3_TREPA	FLAGELLAR FILAMENT 31	6.35e+02
267	4	25.0	1	YGQJ_ECOLI	HYPOTHETICAL 25.5 KD P	6.35e+02	340	4	25.0	286	1	MYOP_XENLA	MYO-INOSITOL-1(OR 4)-M	6.35e+02
268	4	25.0	1	Y529_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	341	4	25.0	286	1	YBL2_YEAST	HYPOTHETICAL 32.7 KD P	6.35e+02
269	4	25.0	1	RECR_SVNY3	RECOMBINATION PROTEIN	6.35e+02	342	4	25.0	287	1	Y011_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02
270	4	25.0	1	CAP1_PIG	ADENYL CYCLASE-ASSOC	6.35e+02	343	4	25.0	290	1	Y43A_MYCPN	HYPOTHETICAL LIPOPROTE	6.35e+02
271	4	25.0	1	PRCD_TOBAC	PROTEASOME DELTA CHAIN	6.35e+02	344	4	25.0	290	1	ARY1_MOUSE	ARYLAMINE N-ACETYLTRAN	6.35e+02
272	4	25.0	1	RPSE_CLOAB	RNA POLYMERASE SIGMA-E	6.35e+02	345	4	25.0	290	1	Y070_NPVAC	HYPOTHETICAL 34.4 KD P	6.35e+02
273	4	25.0	1	YMI7_YEAST	HYPOTHETICAL 27.9 KD P	6.35e+02	346	4	25.0	292	1	IBF3_RAT	INSULIN-LIKE GROWTH FA	6.35e+02
274	4	25.0	1	IAPL_ASFM2	IAP-LIKE PROTEIN P27.	6.35e+02	347	4	25.0	293	1	Y326_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02
275	4	25.0	1	GFP_AEOVI	GREEN FLUORESCENT PROT	6.35e+02	348	4	25.0	293	1	YHM2_YEAST	HYPOTHETICAL 32.2 KD P	6.35e+02
276	4	25.0	1	YG59_YEAST	HYPOTHETICAL 26.7 KD P	6.35e+02	349	4	25.0	293	1	ADH1_ZEALU	ALCOHOL DEHYDROGENASE	6.35e+02
277	4	25.0	1	YC37_CVACA	PROBABLE TRANSCRIPTION	6.35e+02	350	4	25.0	294	1	Y301_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
278	4	25.0	1	Y293_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	351	4	25.0	296	1	ACCD_HAEIN	ACETYL-COENZYME A CARB	6.35e+02
279	4	25.0	1	Y513_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	352	4	25.0	297	1	Y315_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02
280	4	25.0	1	YF45_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	353	4	25.0	297	1	Y1M4_YEAST	HYPOTHETICAL OXIDOREDU	6.35e+02
281	4	25.0	1	LEG3_RABIT	GALACTIN-3 (GALACTOSE-	6.35e+02	354	4	25.0	297	1	GD44_WHEAT	ALPHA/BETA-GLIADIN PRE	6.35e+02
282	4	25.0	1	VP28_YEAST	VACUOLIN-3 (GALACTOSE-	6.35e+02	355	4	25.0	298	1	CC28_YEAST	CELL DIVISION CONTROL	6.35e+02
283	4	25.0	1	Y004_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	356	4	25.0	299	1	ICEL_SPOFR	CASPASE-1 PRECURSOR (E	6.35e+02
284	4	25.0	1	LECA_GRI51	LECTIN IV (GS4) .	6.35e+02	357	4	25.0	302	1	YALD_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02
285	4	25.0	1	GNTR_BACSU	GLUCONATE OPERON TRANS	6.35e+02	358	4	25.0	303	1	YRD2_CAEEL	HYPOTHETICAL 34.4 KD P	6.35e+02
286	4	25.0	1	CYSH_SALTU	PHOSPHODENOSINE PHOSP	6.35e+02	359	4	25.0	303	1	MTM1_METJA	MODIFICATION METHYLASE	6.35e+02
287	4	25.0	1	Y384_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	360	4	25.0	303	1	SFR1_ARATH	PRE-MRNA SPLICING FACT	6.35e+02
288	4	25.0	1	YENR_YEREN	YENR REGULATORY PROTEI	6.35e+02	361	4	25.0	306	1	YEDA_ECOLI	HYPOTHETICAL 32.2 KD P	6.35e+02
289	4	25.0	1	YS21_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02	362	4	25.0	307	1	ETFA_FARDE	ELECTRON TRANSFER FLAV	6.35e+02
290	4	25.0	1	YBEC_ECOLI	HYPOTHETICAL 26.4 KD P	6.35e+02	363	4	25.0	308	1	YBIN_ECOLI	HYPOTHETICAL 34.2 KD P	6.35e+02
291	4	25.0	1	YBEC_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	364	4	25.0	308	1	Y209_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02
292	4	25.0	1	RHIR_RHILV	RHIR REGULATORY PROTEI	6.35e+02	365	4	25.0	308	1	VG53_HSV11	HYPOTHETICAL GENE 53 P	6.35e+02
293	4	25.0	1	YF82_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	366	4	25.0	311	1	LECI_HUMAN	ASIALOGLYCOPROTEIN REC	6.35e+02
294	4	25.0	1	PRG9_SPIOI	PROTEASOME 27 KD SUBUN	6.35e+02	367	4	25.0	312	1	YDIR_ECOLI	PUTATIVE ELECTRON TRAN	6.35e+02
295	4	25.0	1	YMW5_CAEEL	HYPOTHETICAL 27.9 KD P	6.35e+02	368	4	25.0	313	1	FIXB_ECOLI	FIXB PROTEIN.	6.35e+02
296	4	25.0	1	Y44D_MYCPN	HYPOTHETICAL LIPOPROTE	6.35e+02	369	4	25.0	314	1	OLF5_RAT	OLFACTORY RECEPTOR-LIK	6.35e+02
297	4	25.0	1	LACK_SPTMU	LACTOSE PHOSPHOTRANSFE	6.35e+02	370	4	25.0	314	1	YMF7_YEAST	HYPOTHETICAL 34.2 KD P	6.35e+02
298	4	25.0	1	GPX_ECOLI	PHOSPHOGLYCOLATE PHOSP	6.35e+02	371	4	25.0	314	1	GMS1_SCHPO	UDP-GALACTOSE TRANSPOR	6.35e+02
299	4	25.0	1	YC37_CVACA	PROBABLE TRANSCRIPTION	6.35e+02	372	4	25.0	316	1	SRG8_CAEEL	SRG-8 PROTEIN.	6.35e+02
300	4	25.0	1	FECE_ECOLI	IRON(III) DICITRATE TR	6.35e+02	373	4	25.0	316	1	ETFA_CAEEL	PROBABLE ELECTRON TRAN	6.35e+02
301	4	25.0	1	YBUB_BACCI	HYPOTHETICAL PROTEIN I	6.35e+02	374	4	25.0	317	1	Y871_METJA	HYPOTHETICAL PROTEIN B	6.35e+02
302	4	25.0	1	YK60_YEAST	HYPOTHETICAL 29.2 KD P	6.35e+02	375	4	25.0	317	1	YS11_BORBU	HYPOTHETICAL PROTEIN M	6.35e+02
303	4	25.0	1	RL3_MYCGE	SOS RIBOSOMAL PROTEIN	6.35e+02	376	4	25.0	317	1	YM32_PYRHO	HYPOTHETICAL PROTEIN P	6.35e+02
304	4	25.0	1	UPKA_HUMAN	UROPLAKIN TA (UPIA) (U	6.35e+02	377	4	25.0	320	1	Y678_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
305	4	25.0	1	LYCV_BP2ZA	LYSOZYME (EC 3.2.1.17)	6.35e+02	378	4	25.0	320	1	Y054_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02
306	4	25.0	1	MTW2_MORBO	MODIFICATION METHYLASE	6.35e+02	379	4	25.0	320	1	YB9K_YEAST	HYPOTHETICAL 36.0 KD P	6.35e+02
307	4	25.0	1	NMA_HUMAN	PUTATIVE TRANSMEMBRANE	6.35e+02	380	4	25.0	321	1	MOAA_HELPY	MOLYBDENUM COFACTOR B1	6.35e+02
308	4	25.0	1	VP31_FRG3V	EARLY 31 KD PROTEIN.	6.35e+02	381	4	25.0	322	1	YK02_YEAST	HYPOTHETICAL 36.6 KD P	6.35e+02
309	4	25.0	1	Y725_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	382	4	25.0	323	1	CATV_NPVAC	VIRAL CATHEPSIN (EC 3.	6.35e+02
310	4	25.0	1	YXBG_BACSU	HYPOTHETICAL OXIDOREDU	6.35e+02	383	4	25.0	323	1	HEM2_STAUA	DELTA-AMINOLEVULINIC A	6.35e+02
311	4	25.0	1	PENK_XENLA	PROENKEPHALIN AL1 PRECU	6.35e+02	384	4	25.0	324	1	Y700_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
312	4	25.0	1	LYCV_BP03	LYSOZYME (EC 3.2.1.17)	6.35e+02	385	4	25.0	324	1	CATV_NPVOP	VIRAL CATHEPSIN (EC 3.	6.35e+02
313	4	25.0	1	DIVB_BACSU	DIVISION INITIATION PR	6.35e+02	386	4	25.0	325	1	ETFA_BACSU	ELECTRON TRANSFER FLAV	6.35e+02
314	4	25.0	1	NUC_SERMA	NUCLEASE PRECURSOR (EC	6.35e+02	387	4	25.0	325	1	CYL_HUMAN	CYTOCHROME C1, HEME PR	6.35e+02
315	4	25.0	1	LY4E_MOUSE	T-CELL SURFACE GLYCOPR	6.35e+02	388	4	25.0	325	1	CA19_RAT	COLLAGEN ALPHA 1(IX) C	6.35e+02

389	4	25.0	1	YTG3_YEAST	HYPOTHETICAL 36.1 KD P	6.35e+02	462	4	25.0	375	1	YHCV_ECOLI	HYPOTHETICAL 43.1 KD P	6.35e+00
390	4	25.0	1	VMP_CAMV5	MOVEMENT PROTEIN (CELL)	6.35e+02	463	4	25.0	375	1	ADH2_ORISA	ALCOHOL DEHYDROGENASE	6.35e+02
391	4	25.0	1	VMP_CAMV6	MOVEMENT PROTEIN (CELL)	6.35e+02	464	4	25.0	375	1	ADH1_RANPE	ALCOHOL DEHYDROGENASE	6.35e+02
392	4	25.0	1	VMP_CAMV7	MOVEMENT PROTEIN (CELL)	6.35e+02	465	4	25.0	375	1	ADHA_UROHA	ALCOHOL DEHYDROGENASE	6.35e+02
393	4	25.0	1	VMP_CAMV8	MOVEMENT PROTEIN (CELL)	6.35e+02	466	4	25.0	375	1	ADH1_CHICK	ALCOHOL DEHYDROGENASE	6.35e+02
394	4	25.0	1	VMP_CAMV9	MOVEMENT PROTEIN (CELL)	6.35e+02	467	4	25.0	375	1	ADHA_RAT	ALCOHOL DEHYDROGENASE	6.35e+02
395	4	25.0	1	VMP_CAMV10	MOVEMENT PROTEIN (CELL)	6.35e+02	468	4	25.0	377	1	YGLT_YEAST	HYPOTHETICAL 42.5 KD P	6.35e+02
396	4	25.0	1	GPT_SULAC	PUTATIVE UDP-N-ACETYLGL	6.35e+02	469	4	25.0	377	1	PGT1_RAT	GERANYLGERANYL TRANSFER	6.35e+02
397	4	25.0	1	YH13_LACILA	HYPOTHETICAL 38.0 KD P	6.35e+02	470	4	25.0	378	1	YK84_CAEEL	HYPOTHETICAL 43.0 KD P	6.35e+02
398	4	25.0	1	YCG1_YEAST	HYPOTHETICAL 38.3 KD P	6.35e+02	471	4	25.0	378	1	ADH4_PEA	ALCOHOL DEHYDROGENASE	6.35e+02
399	4	25.0	1	RBSR_ECOLI	RIBOSE OPERON REPRESSO	6.35e+02	472	4	25.0	379	1	Y024_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02
400	4	25.0	1	IPNS_STRUJ	ISOPENTICILLIN N SYNTHASE	6.35e+02	473	4	25.0	379	1	ADH2_STRCA	ALCOHOL DEHYDROGENASE	6.35e+02
401	4	25.0	1	LDHX_HUMAN	L-LACTATE DEHYDROGENAS	6.35e+02	474	4	25.0	379	1	ADH1_ARATH	ALCOHOL DEHYDROGENASE	6.35e+02
402	4	25.0	1	LDHX_RAT	L-LACTATE DEHYDROGENAS	6.35e+02	475	4	25.0	380	1	ADH3_SOLTU	ALCOHOL DEHYDROGENASE	6.35e+02
403	4	25.0	1	LDHM_MOUSE	L-LACTATE DEHYDROGENAS	6.35e+02	476	4	25.0	380	1	ADH2_SOLTU	ALCOHOL DEHYDROGENASE	6.35e+02
404	4	25.0	1	LDHM_BOVIN	L-LACTATE DEHYDROGENAS	6.35e+02	477	4	25.0	380	1	RPOA_CHLTR	RNA-DIRECTED RNA POLYM	6.35e+02
405	4	25.0	1	LDHM_PIG	L-LACTATE DEHYDROGENAS	6.35e+02	478	4	25.0	380	1	PLVA_COLGL	PECTIN LYASE PRECURSOR	6.35e+02
406	4	25.0	1	YGHM_YEAST	HYPOTHETICAL 37.1 KD P	6.35e+02	479	4	25.0	381	1	YE21_CAEEL	HYPOTHETICAL 42.2 KD P	6.35e+02
407	4	25.0	1	LDHM_CHICK	L-LACTATE DEHYDROGENAS	6.35e+02	480	4	25.0	382	1	YJ07_YEAST	HYPOTHETICAL 44.9 KD P	6.35e+02
408	4	25.0	1	IPAD_SHIDI	37 KD MEMBRANE ANTIGEN	6.35e+02	481	4	25.0	385	1	YH10_ECOLI	HYPOTHETICAL 41.2 KD L	6.35e+02
409	4	25.0	1	LDH_DROME	L-LACTATE DEHYDROGENAS	6.35e+02	482	4	25.0	385	1	Y001_MYCTU	HYPOTHETICAL 39.8 KD P	6.35e+02
410	4	25.0	1	LDHM_ANAPL	L-LACTATE DEHYDROGENAS	6.35e+02	483	4	25.0	385	1	OM23_HAEIN	OUTER MEMBRANE PROTEIN	6.35e+02
411	4	25.0	1	LDHM_RAT	L-LACTATE DEHYDROGENAS	6.35e+02	484	4	25.0	386	1	ADT2_SOLTU	ADP ATP CARRIER PROTEIN	6.35e+02
412	4	25.0	1	Y009_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02	485	4	25.0	386	1	OM24_HAEIN	OUTER MEMBRANE PROTEIN	6.35e+02
413	4	25.0	1	UL16_HSV6U	PROTEIN U65	6.35e+02	486	4	25.0	388	1	SAOX_AETSP	SARCOSINE OXIDASE (EC	6.35e+02
414	4	25.0	1	YTCB_PSEB	HYPOTHETICAL 35.9 KD P	6.35e+02	487	4	25.0	388	1	NHAA_ECOLI	NA(+)/H(+) ANTIporter	6.35e+02
415	4	25.0	1	ETFA_CLOAB	ELECTRON TRANSFER FLAV	6.35e+02	488	4	25.0	388	1	YFHS_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02
416	4	25.0	1	Y941_METUA	HYPOTHETICAL PROTEIN M	6.35e+02	489	4	25.0	388	1	RALL_ONCVO	RAL-1 PROTEIN PRECURSO	6.35e+02
417	4													

535	4	25.0	425	1	KDTA_ECOLI	3-DEOXY-D-MANNO-OCTULO	6.35e+02	608	4	25.0	492	1	CG21_CANAL	G2/MITOTIC-SPECIFIC CY	6.35e+02
536	4	25.0	426	1	WCAK_ECOLI	COLANIC ACID BIOSYNTHET	6.35e+02	609	4	25.0	493	1	HEMZ_ORISA	FERROCHELATASE PRECURS	6.35e+02
537	4	25.0	427	1	Y420_VARV	PROTEIN A20.	6.35e+02	610	4	25.0	494	1	ACH3_BOVIN	NEURONAL ACETYLCHOLINE	6.35e+02
538	4	25.0	428	1	Y4FK_RHISN	PUTATIVE TRANSCRIPTION	6.35e+02	611	4	25.0	495	1	YAOI_SCHPO	PUTATIVE TRANSPORTER C	6.35e+02
539	4	25.0	429	1	GLFD_ECOLI	GLYCOPHYTHICAL 46.4 KD P	6.35e+02	612	4	25.0	496	1	PSG4_YEAST	PREGNANCY-SPECIFIC BET	6.35e+02
540	4	25.0	430	1	GLGC_ANASP	GLUCOSE-1-PHOSPHATE AD	6.35e+02	613	4	25.0	497	1	HKXG_YEAST	GLUCOKINASE (EC 2.7.1.	6.35e+02
541	4	25.0	431	1	ACDL_PIG	ACYL-COA DEHYDROGENASE	6.35e+02	614	4	25.0	500	1	YH92_CAEEL	HYPOTHETICAL 54.5 KD T	6.35e+02
542	4	25.0	432	1	ACDL_RAT	ACYL-COA DEHYDROGENASE	6.35e+02	615	4	25.0	501	1	MYSU_RABIT	MYOSIN HEAVY CHAIN, EM	6.35e+02
543	4	25.0	433	1	IF5_HUMAN	EUKARYOTIC TRANSLATION	6.35e+02	616	4	25.0	502	1	HST1_YEAST	HST1 PROTEIN (HOMOLOGO	6.35e+02
544	4	25.0	434	1	UL43_HSV11	HYPOTHETICAL PROTEIN M	6.35e+02	617	4	25.0	503	1	Y038_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02
545	4	25.0	435	1	UDG_RICPR	MEMBRANE PROTEIN UL43	6.35e+02	618	4	25.0	504	1	Y11R_YEAST	HYPOTHETICAL 58.3 KD P	6.35e+02
546	4	25.0	436	1	PHSA_PHAVU	UDP-GLUCOSE 6-DEHYDROG	6.35e+02	619	4	25.0	505	1	YG2M_YEAST	HYPOTHETICAL 58.2 KD P	6.35e+02
547	4	25.0	437	1	ZN10_HUMAN	PHASCOLIN, ALPHA-TYPE	6.35e+02	620	4	25.0	506	1	Y026_CAEEL	HYPOTHETICAL 56.9 KD P	6.35e+02
548	4	25.0	438	1	ZN10_HUMAN	ZINC FINGER PROTEIN 10	6.35e+02	621	4	25.0	507	1	ATP0_ARATH	ATP SYNTHASE ALPHA CHA	6.35e+02
549	4	25.0	439	1	GLNA_THEMEA	GLUTAMINE SYNTHETASE (	6.35e+02	622	4	25.0	508	1	Y1K4_YEAST	HYPOTHETICAL 59.2 KD P	6.35e+02
550	4	25.0	440	1	YLA4_SCHPO	HYPOTHETICAL 51.9 KD P	6.35e+02	623	4	25.0	509	1	HSK1_SCHPO	CELL CYCLE PROTEIN KIN	6.35e+02
551	4	25.0	441	1	HSTC_CAEBR	HEAT SHOCK 70 KD PROTE	6.35e+02	624	4	25.0	510	1	ATP0_SOYBN	ATP SYNTHASE ALPHA CHA	6.35e+02
552	4	25.0	442	1	ETBR_RAT	ENDOTHELIN B RECEPTOR	6.35e+02	625	4	25.0	511	1	YDEX_ECOLI	HYPOTHETICAL ABC TRANS	6.35e+02
553	4	25.0	443	1	Y1H1_YEAST	HYPOTHETICAL 51.3 KD P	6.35e+02	626	4	25.0	512	1	MDOG_ECOLI	HYPOTHETICAL ABC TRANS	6.35e+02
554	4	25.0	444	1	YCAJ_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	627	4	25.0	513	1	HM22_ARATH	PERIPLASMIC GLUCANS B1	6.35e+02
555	4	25.0	445	1	PRSA_PSESY	MRSA PROTEIN HOMOLOG.	6.35e+02	628	4	25.0	514	1	K2C5_XENLA	FERROCHELATASE, CHLORO	6.35e+02
556	4	25.0	446	1	PEPC_LACHE	AMINOPEPTIDASE C (EC 3	6.35e+02	629	4	25.0	515	1	R1T1_YEAST	KERATIN, TYPE II CYTOS	6.35e+02
557	4	25.0	447	1	DNA4_MYCCA	CHROMOSOMAL REPLICATIO	6.35e+02	630	4	25.0	516	1	K2C5_XENLA	TRNA A64-2'-O-RIBOSYL	6.35e+02
558	4	25.0	448	1	ENVZ_SALTI	OSMOLARITY SENSOR PROT	6.35e+02	631	4	25.0	517	1	HEN0_CHICK	5-AMINOLEVULINIC ACID	6.35e+02
559	4	25.0	449	1	ENVZ_ECOLI	OSMOLARITY SENSOR PROT	6.35e+02	632	4	25.0	518	1	HEN0_CHICK	FERROCHELATASE PRECURS	6.35e+02
560	4	25.0	450	1	ENVZ_SALTY	OSMOLARITY SENSOR PROT	6.35e+02	633	4	25.0	519	1	HEMZ_CUCSA	FERROCHELATASE PRECURS	6.35e+02
561	4	25.0	451	1	YU10_HSV7J	U10 PROTEIN.	6.35e+02	634	4	25.0	520	1	LIM_HALRO	HOMOBBOX PROTEIN LIM (	6.35e+02
562	4	25.0	452	1	IF5_MAIZE	EUKARYOTIC TRANSLATION	6.35e+02	635	4	25.0	521	1	TUPL_CANAL	SPINDLE POLE BODY ASSO	6.35e+02
563	4	25.0	453	1	EUTB_SALTY	ETHANOLAMINE AMMONIA-L	6.35e+02	636	4	25.0	522	1	SAD1_SCHPO	INTERMEDIATE FILAMENT	6.35e+02
564	4	25.0	454	1	EUTB_ECOLI	ETHANOLAMINE AMMONIA-L	6.35e+02	637	4	25.0	523	1	YMB0_YEAST	HYPOHETICAL 57.7 KD P	6.35e+02
565	4	25.0	455	1	Y003_BORBU	HYPOTHETICAL PROTEIN B	6.35e+02	638	4	25.0	524	1	YMB0_YEAST	HYPOHETICAL 57.7 KD P	6.35e+02
566	4	25.0	456	1	Y044_SCHPO	HYPOTHETICAL 51.4 KD P	6.35e+02	639	4	25.0	525	1	YMB0_YEAST	HYPOHETICAL 57.7 KD P	6.35e+02
567	4	25.0	457	1	YHCL_ECOLI	HYPOTHETICAL 48.8 KD P	6.35e+02	640	4	25.0	526	1	YMB0_YEAST	HYPOHETICAL 57.7 KD P	6.35e+02
568	4	25.0	458	1	GCAD_BACSU	UDP-N-ACETYLGLUCOSAMIN	6.35e+02	641	4	25.0	527	1	ION3_CARAU	PUTATIVE 2-ISOPROPYLMA	6.35e+02
569	4	25.0	459	1	VCCL_PEA	PROVINCILIN PRECURSOR.	6.35e+02	642	4	25.0	528	1	GLGS_SOLTU	GLUCOSE-1-PHOSPHATE AD	6.35e+02
570	4	25.0	460	1	Y454_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	643	4	25.0	529	1	GAG_HV2KR	GAG POLYPROTEIN [CONTA	6.35e+02
571	4	25.0	461	1	YCBJ_ECOLI	HYPOTHETICAL 50.9 KD T	6.35e+02	644	4	25.0	530	1	YHAG_ECOLI	HYPOHETICAL 56.4 KD P	6.35e+02
572	4	25.0	462	1	YFHD_HAEIN	HYPOTHETICAL PROTEIN H	6.35e+02	645	4	25.0	531	1	K2C4_MOUSE	KERATIN, TYPE II CYTOS	6.35e+02
573	4	25.0	463	1	COGM_MOUSE	MACROPHAGE METALLOELAS	6.35e+02	646	4	25.0	532	1	YBY8_YEAST	HYPOHETICAL 60.5 KD P	6.35e+02
574	4	25.0	464	1	YH06_YEAST	HYPOHETICAL 51.1 KD P	6.35e+02	647	4	25.0	533	1	YMB4_YEAST	HYPOHETICAL 56.8 KD P	6.35e+02
575	4	25.0	465	1	YF22_YEAST	PUTATIVE 62.1 KD TRANS	6.35e+02	648	4	25.0	534	1	YMB4_YEAST	HYPOHETICAL 56.8 KD P	6.35e+02
576	4	25.0	466	1	GAC1_RAT	GAMMA-AMINOBUTYRIC-ACI	6.35e+02	649	4	25.0	535	1	FM05_MOUSE	DIMETHYLANILINE MONOOX	6.35e+02
577	4	25.0	467	1	GAC2_HUMAN	GAMMA-AMINOBUTYRIC-ACI	6.35e+02	650	4	25.0	536	1	REF2_YEAST	RNA END RECOMBINATION	6.35e+02
578	4	25.0	468	1	NRAM_TARIS	NEURAMINIDASE (EC 3.2.	6.35e+02	651	4	25.0	537	1	YAJE_SCHPO	HYPOHETICAL 59.0 KD P	6.35e+02
579	4	25.0	469	1	ATPB_PECFR	ATP SYNTHASE BETA CHAI	6.35e+02	652	4	25.0	538	1	Y255_METTH	O-SIALOGLYCOPROTEIN EN	6.35e+02
580	4	25.0	470	1	Y123_MYCCE	HYPOTHETICAL PROTEIN M	6.35e+02	653	4	25.0	539	1	K2C4_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02
581	4	25.0	471	1	PHR_ECOLI	DEOXYRIBODIPYRIMIDINE	6.35e+02	654	4	25.0	540	1	YMB4_YEAST	HYPOHETICAL 56.8 KD P	6.35e+02
582	4	25.0	472	1	COG2_HORSE	COLLAGENASE 3 PRECURSO	6.35e+02	655	4	25.0	541	1	YMB4_YEAST	HYPOHETICAL 56.8 KD P	6.35e+02
583	4	25.0	473	1	TRAL_CLOPE	TRANSPOSASE FOR INSERT	6.35e+02	656	4	25.0	542	1	YMB4_YEAST	HYPOHETICAL 56.8 KD P	6.35e+02
584	4	25.0	474	1	GSHB_XENLA	GLUTATHIONE SYNTHETASE	6.35e+02	657	4	25.0	543	1	YIEC_ECOLI	POTENTIAL OUTER MEMBRA	6.35e+02
585	4	25.0	475	1	L1PL_MOUSE	LIPOPROTEIN LIPASE PRE	6.35e+02	658	4	25.0	544	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
586	4	25.0	476	1	NOEK_RHISN	PHOSPHOMANNOMUTASE (EC	6.35e+02	659	4	25.0	545	1	TCPE_CAEEL	HEXOSE TRANSPORTER HAT	6.35e+02
587	4	25.0	477	1	GAC2_CHICK	GAMMA-AMINOBUTYRIC-ACI	6.35e+02	660	4	25.0	546	1	TCPE_CAEEL	T-COMPLEX PROTEIN 1, E	6.35e+02
588	4	25.0	478	1	Y123_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	661	4	25.0	547	1	IMAI_SCHPO	IMPORTIN ALPHA SUBUNIT	6.35e+02
589	4	25.0	479	1	L1PL_MOUSE	LIPOPROTEIN LIPASE PRE	6.35e+02	662	4	25.0	548	1	CP1B_HUMAN	CYTOCHROME P450 1B1 (E	6.35e+02
590	4	25.0	480	1	L1PL_PAPAN	LIPOPROTEIN LIPASE PRE	6.35e+02	663	4	25.0	549	1	DPPE_BACSU	DIPEPTIDE-BINDING PROT	6.35e+02
591	4	25.0	481	1	GAC2_BOVIN	GAMMA-AMINOBUTYRIC-ACI	6.35e+02	664	4	25.0	550	1	TCPG_HUMAN	T-COMPLEX PROTEIN 1, G	6.35e+02
592	4	25.0	482	1	OAR_HELVI	OCTOPAMINE RECEPTOR.	6.35e+02	665	4	25.0	551	1	TCPG_HUMAN	T-COMPLEX PROTEIN 1, G	6.35e+02
593	4	25.0	483	1	PR4_SCHPO	SERINE/THREONINE-PROTE	6.35e+02	666	4	25.0	552	1	TCPG_HUMAN	T-COMPLEX PROTEIN 1, G	6.35e+02
594	4	25.0	484	1	L1PL_SHEEP	LIPOPROTEIN LIPASE PRE	6.35e+02	667	4	25.0	553	1	TCPG_HUMAN	T-COMPLEX PROTEIN 1, G	6.35e+02
595	4	25.0	485	1	IC68_HUMAN	CASPASE-8 PRECURSOR (E	6.35e+02	668	4	25.0	554	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
596	4	25.0	486	1	Y066_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	669	4	25.0	555	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
597	4	25.0	487	1	HMCM_MOUSE	HYDROXYMETHYLGLUTARYL-	6.35e+02	670	4	25.0	556	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
598	4	25.0	488	1	IMDH_HELPY	INOSINE-5'-MONOPHOSPHA	6.35e+02	671	4	25.0	557	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
599	4	25.0	489	1	K2C8_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02	672	4	25.0	558	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
600	4	25.0	490	1	Y142_NPVOP	HYPOTHETICAL 55.8 KD P	6.35e+02	673	4	25.0	559	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
601	4	25.0	491	1	CATA_PROMI	CATALASE (EC 1.11.1.6)	6.35e+02	674	4	25.0	560	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
602	4	25.0	492	1	YUP2_CAEEL	HYPOTHETICAL 54.5 KD P	6.35e+02	675	4	25.0	561	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
603	4	25.0	493	1	HMCS_YEAST	HYDROXYMETHYLGLUTARYL-	6.35e+02	676	4	25.0	562	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
604	4	25.0	494	1	YF66_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	677	4	25.0	563	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
605	4	25.0	495	1	GAL7_STRMU	GALACTOSE-1-PHOSPHATE	6.35e+02	678	4	25.0	564	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
606	4	25.0	496	1	YNN4_YEAST	HYPOTHETICAL 54.9 KD P	6.35e+02	679	4	25.0	565	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02
607	4	25.0	497	1	PEPD_HUMAN	XAA-PRO DIPEPTIDASE (E	6.35e+02	680	4	25.0	566	1	YH1J_ECOLI	HYPOHETICAL 61.2 KD P	6.35e+02

681	4	25.0	560	1	TATR_NPVCF	TRANS-ACTIVATING TRANS	6.35e+02	754	1	659	1	Y05L_MYCLE	HYPOTHETICAL 71.2 KD P	6.35e+02
682	4	25.0	561	1	DIM_ARATH	CELL ELONGATION PROTEIN	6.35e+02	755	4	659	1	YH5_YEAST	HYPOTHETICAL 75.9 KD P	6.35e+02
683	4	25.0	561	1	CEAC_ECOLI	CLORACIN (EC 3.1.-.-) (	6.35e+02	756	4	663	1	GR72_MAIZE	IMMUNOGLOBULIN BINDING	6.35e+02
684	4	25.0	562	1	LCB2_HUMAN	SERINE PALMITOYLTRANSF	6.35e+02	757	4	663	1	GR73_MAIZE	IMMUNOGLOBULIN BINDING	6.35e+02
685	4	25.0	563	1	K2CE_KLALA	KERATIN, TYPE II CYTOS	6.35e+02	758	4	664	1	X366_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02
686	4	25.0	563	1	K2CB_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02	759	4	666	1	Y079_SCHPO	HYPOTHETICAL 74.2 KD P	6.35e+02
687	4	25.0	566	1	YJ19_SUNY3	HYPOTHETICAL 65.1 KD P	6.35e+02	760	4	670	1	CATA_PENJA	CATALASE (EC 1.11.1.6)	6.35e+02
688	4	25.0	569	1	NTR_MAIZE	FERREDOXIN--NITRITE RE	6.35e+02	761	4	671	1	KRP2_RAT	KINESIN-RELATED PROTEI	6.35e+02
689	4	25.0	571	1	YB63_SCHPO	HYPOTHETICAL 63.9 KD P	6.35e+02	762	4	676	1	UL06_HSV1	VIRION PROTEIN UL6	6.35e+02
690	4	25.0	572	1	YH4_YEAST	HEAT SHOCK PROTEIN 70	6.35e+02	763	4	682	1	TUPL_KLULA	TRANSCRIPTIONAL REPRES	6.35e+02
691	4	25.0	572	1	MGIA_MYCPN	GALACTOSIDE TRANSPORT	6.35e+02	764	4	683	1	YJCQ_ECOLI	HYPOTHETICAL 76.1 KD P	6.35e+02
692	4	25.0	575	1	FLAA_CAMJE	FLAGELLIN A.	6.35e+02	765	4	686	1	YPR4_CAEEL	HYPOTHETICAL 77.0 KD T	6.35e+02
693	4	25.0	575	1	FLAB_CAMJE	FLAGELLIN B.	6.35e+02	766	4	686	1	TP6B_METJA	TYPE II DNA TOPOISOMER	6.35e+02
694	4	25.0	576	1	YI10_YEAST	HYPOTHETICAL 65.7 KD P	6.35e+02	767	4	686	1	TRFE_ANAPL	OVOTRANSFERRIN.	6.35e+02
695	4	25.0	580	1	YH8X_YEAST	HYPOTHETICAL 68.1 KD P	6.35e+02	768	4	688	1	THIB_XENLA	THYROID HORMONE--INDUCE	6.35e+02
696	4	25.0	583	1	EDD_ZIFMO	PHOSPHOGLUCONATE DEHYD	6.35e+02	769	4	688	1	YEQ7_YEAST	HYPOTHETICAL 79.5 KD P	6.35e+02
697	4	25.0	584	1	COAT_PAVC2	COAT PROTEIN VP2.	6.35e+02	770	4	688	1	LEP_STAEP	LIPASE PRECURSOR (EC 3	6.35e+02
698	4	25.0	585	1	KPK_BACLI	PYRUVATE KINASE (EC 2.	6.35e+02	771	4	695	1	GFAL_SCHPO	PUTATIVE GLUCOSAMINE--	6.35e+02
699	4	25.0	588	1	DECA_DROME	DECAPENTAPLEGIC PROTEI	6.35e+02	772	4	695	1	PIP6_BOVIN	1-PHOSPHATIDYLINOSITOL	6.35e+02
700	4	25.0	591	1	GAG_MMTVB	GAG POLYPROTEIN [CONTA	6.35e+02	773	4	699	1	HS83_ARATH	HEAT SHOCK PROTEIN 81-	6.35e+02
701	4	25.0	591	1	GAG_MMTVC	GAG POLYPROTEIN [CONTA	6.35e+02	774	4	700	1	FLHA_CAUCR	FLAGELLAR BIOSYNTHESIS	6.35e+02
702	4	25.0	592	1	OPF2_DICDI	DIHYDROLIPOAMIDE ACETY	6.35e+02	775	4	702	1	MCAR_CRIGR	MITOTIC CENTROMERE-ASS	6.35e+02
703	4	25.0	593	1	COR1_SCHPO	MITOSIS INDUCER PROTEI	6.35e+02	776	4	703	1	UVRD_MYCGE	PUTATIVE DNA HELICASE	6.35e+02
704	4	25.0	594	1	UL93_HCMVA	PROTEIN UL93.	6.35e+02	777	4	707	1	KPC2_CAEEL	PROTEIN KINASE C-LIKE	6.35e+02
705	4	25.0	594	1	NIR_SPIOL	FERREDOXIN--NITRITE RE	6.35e+02	778	4	711	1	OMC_NEIGO	OUTER MEMBRANE PROTEIN	6.35e+02
706	4	25.0	594	1	DBP9_YEAST	PROBABLE ATP-DEPENDENT	6.35e+02	779	4	711	1	TUPL_YEAST	GLUCOSE REPRESSION REG	6.35e+02
707	4	25.0	599	1	AFAM_HUMAN	AFAMIN PRECURSOR (ALPH	6.35e+02	780	4	715	1	HS82_MAIZE	HEAT SHOCK PROTEIN 82.	6.35e+02
708	4	25.0	599	1	CAPD_STAAR	CAPD PROTEIN.	6.35e+02	781	4	716	1	KIF2_MOUSE	KINESIN-LIKE PROTEIN K	6.35e+02
709	4	25.0	600	1	XOXF_PARDE	PUTATIVE DEHYDROGENASE	6.35e+02	782	4	724	1	SP90_MOUSE	PRESYNAPTIC DENSITY PR	6.35e+02
710	4	25.0	601	1	Y98_SCHPO	HYPOTHETICAL 70.6 KD P	6.35e+02	783	4	727	1	COAT_FPV	COAT PROTEIN VP1 [CONT	6.35e+02
711	4	25.0	605	1	VCIA_GOSHI	VICILIN PRECURSOR (ALP	6.35e+02	784	4	728	1	ARP_ECOLI	ANKRYIN-LIKE REGULATOR	6.35e+02
712	4	25.0	607	1	YHML_YEAST	HYPOTHETICAL 69.8 KD P	6.35e+02	785	4	729	1	NTT7_RAT	SODIUM- AND CHLORIDE-D	6.35e+02
713	4	25.0	607	1	YHVC_YEAST	PROBABLE ARGINYL--TRNA	6.35e+02	786	4	729	1	FHUE_ECOLI	OUTER-MEMBRANE RECEPT	6.35e+02
714	4	25.0	608	1	YD6C_SCHPO	HYPOTHETICAL 67.5 KD P	6.35e+02	787	4	732	1	YH8K_ECOLI	HYPOTHETICAL 83.4 KD P	6.35e+02
715	4	25.0	608	1	GHR_CHICK	GROWTH HORMONE RECEPT	6.35e+02	788	4	732	1	YAGR_ECOLI	LDLC PROTEIN.	6.35e+02
716	4	25.0	609	1	YHML_YEAST	SPORE-SPECIFIC PROTEIN	6.35e+02	789	4	738	1	LDLC_HUMAN	LDLC PROTEIN.	6.35e+02
717	4	25.0	611	1	IAP_CHICK	INHIBITOR OF APOPTOSIS	6.35e+02	790	4	738	1	PECL_HUMAN	PLATELET ENDOTHELIAL C	6.35e+02
718	4	25.0	614	1	YA3A_SCHPO	HYPOTHETICAL 67.3 KD T	6.35e+02	791	4	740	1	GNT5_RAT	ALPHA-1,3(6)-MANNOSYL	6.35e+02
719	4	25.0	616	1	YV06_CAEEL	HYPOTHETICAL 68.2 KD P	6.35e+02	792	4	740	1	PECL_PIG	PLATELET ENDOTHELIAL C	6.35e+02
720	4	25.0	617	1	PGFI_RICCO	XYLOPHOSPHATE--FRUCTOS	6.35e+02	793	4	745	1	YJ00_YEAST	HYPOTHETICAL 84.3 KD P	6.35e+02
721	4	25.0	618	1	GGAL_YEAST	GLYCOGEN SYNTHESIS INI	6.35e+02	794	4	746	1	FEP4_ECOLI	FERRITEROBACTIN RECE	6.35e+02
722	4	25.0	620	1	Y241_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02	795	4	752	1	DRS1_YEAST	PUTATIVE ATP-DEPENDENT	6.35e+02
723	4	25.0	620	1	DNAK_PORPU	DNAK PROTEIN (HEAT SHO	6.35e+02	796	4	752	1	METE_ECOLI	5-METHYL-TETRAHYDROPTER	6.35e+02
724	4	25.0	621	1	Y001_CAEEL	HYPOTHETICAL 69.9 KD P	6.35e+02	797	4	754	1	MBPL_KLULA	TRANSCRIPTION FACTOR M	6.35e+02
725	4	25.0	621	1	Y0DA_CAEEL	HYPOTHETICAL 69.5 KD P	6.35e+02	798	4	755	1	COMP_RAT	CARTILAGE OLIGOMERIC M	6.35e+02
726	4	25.0	621	1	SRPR_YEAST	SIGNAL RECOGNITION PAR	6.35e+02	799	4	756	1	Y328_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02
727	4	25.0	621	1	ILVB_MYCAV	PROBABLE ACETOLACTATE	6.35e+02	800	4	756	1	PIP6_HUMAN	1-PHOSPHATIDYLINOSITOL	6.35e+02
728	4	25.0	621	1	HEM1_AGABI	5-AMINOLEVULINIC ACID	6.35e+02	801	4	759	1	Y363_METJA	HYPOTHETICAL MCM-TYPE	6.35e+02
729	4	25.0	622	1	YAE7_YEAST	HYPOTHETICAL 72.1 KD P	6.35e+02	802	4	766	1	DOC2_MOUSE	MITOGEN-RESPONSIVE PHO	6.35e+02
730	4	25.0	623	1	EX70_YEAST	70 KD EXOCYST COMPLEX	6.35e+02	803	4	768	1	SC23_YEAST	PROTEIN TRANSPORT PROT	6.35e+02
731	4	25.0	627	1	MUTL_BACSU	DNA MISMATCH REPAIR PR	6.35e+02	804	4	772	1	TLE3_HUMAN	TRANSDUCCIN-LIKE ENHANC	6.35e+02
732	4	25.0	628	1	LO_HUMAN	LUTHERAN BLOOD GROUP G	6.35e+02	805	4	773	1	MAK5_YEAST	ATP-DEPENDENT RNA HELI	6.35e+02
733	4	25.0	629	1	K2C3_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02	806	4	780	1	Z084_XENLA	OOCYTE ZINC FINGER PRO	6.35e+02
734	4	25.0	630	1	Y242_MYCGE	HYPOTHETICAL PROTEIN M	6.35e+02	807	4	781	1	YG5B_YEAST	HYPOTHETICAL 89.5 KD P	6.35e+02
735	4	25.0	631	1	HMWR_MOUSE	HYALURONAN-MEDIATED MO	6.35e+02	808	4	782	1	YVW2_YEAST	HYPOTHETICAL 88.8 KD P	6.35e+02
736	4	25.0	632	1	Y242_MYCPN	HYPOTHETICAL PROTEIN M	6.35e+02	809	4	782	1	YQ91_CAEEL	HYPOTHETICAL 90.1 KD P	6.35e+02
737	4	25.0	632	1	YK65_CAEEL	HYPOTHETICAL 71.9 KD P	6.35e+02	810	4	784	1	YQ91_CAEEL	GC-RICH SEQUENCE DNA-B	6.35e+02
738	4	25.0	633	1	PABP_XENLA	POLYADENYLATE-BINDING	6.35e+02	811	4	784	1	IF2_ENTIFC	TRANSLATION INITIATION	6.35e+02
739	4	25.0	635	1	TPOR_HUMAN	THROMBOPOIETIN RECEPT	6.35e+02	812	4	793	1	REGA_DICDI	3',5'-CYCLIC-NUCLEOTID	6.35e+02
740	4	25.0	636	1	PAB1_MOUSE	POLYADENYLATE-BINDING	6.35e+02	813	4	800	1	KIN4_YEAST	SERINE/THREONINE-PROTE	6.35e+02
741	4	25.0	636	1	PAB1_HUMAN	POLYADENYLATE-BINDING	6.35e+02	814	4	803	1	GYRB_BUCAP	DNA GYRASE SUBUNIT B (	6.35e+02
742	4	25.0	638	1	K220_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02	815	4	804	1	R1RL_HSV6U	RIBONUCLEOSIDE-DIPHOSPH	6.35e+02
743	4	25.0	639	1	HS74_PARLI	HEAT SHOCK 70 KD PROTE	6.35e+02	816	4	807	1	HYS4_STAAR	HYALURONATE LYASE PREC	6.35e+02
744	4	25.0	642	1	R5D_NEIGO	RNA POLYMERASE SIGMA F	6.35e+02	817	4	814	1	COLA_VIBAL	MICROBIAL COLLAGENASE	6.35e+02
745	4	25.0	643	1	Y903_METJA	HYPOTHETICAL PROTEIN M	6.35e+02	818	4	816	1	YG3A_YEAST	HYPOTHETICAL 92.7 KD P	6.35e+02
746	4	25.0	643	1	K2C1_HUMAN	KERATIN, TYPE II CYTOS	6.35e+02	819	4	823	1	PARC_STRPN	TOPOISOMERASE IV SUBUN	6.35e+02
747	4	25.0	644	1	NFM_RABIT	NEUROFILAMENT TRIPLET	6.35e+02	820	4	833	1	HSF_YEAST	HEAT SHOCK FACTOR PROT	6.35e+02
748	4	25.0	645	1	IME2_YEAST	MEIOSIS INDUCTION PROT	6.35e+02	821	4	838	1	YK59_YEAST	HEAT SHOCK FACTOR 8 KD P	6.35e+02
749	4	25.0	647	1	SKO1_YEAST	CRE-BINDING 82IP PROTE	6.35e+02	822	4	842	1	IF2_TREPA	TRANSLATION INITIATION	6.35e+02
750	4	25.0	651	1	HS70_ONCMY	HEAT SHOCK COGNATE 70	6.35e+02	823	4	843	1	DPO1_HPBT	DNA POLYMERASE (EC 2.7	6.35e+02
751	4	25.0	653	1	Y07S_MYCTU	HYPOTHETICAL 69.5 KD P	6.35e+02	824	4	845	1	NFM_RAT	NEUROFILAMENT TRIPLET	6.35e+02
752	4	25.0	654	1	ICAL_RAT	CALPAIN INHIBITOR (CAL	6.35e+02	825	4	847	1	HEX_VIBU	BETA-HEXOSAMINIDASE (E	6.35e+02
753	4	25.0	655	1	MA12_MOUSE	MANNOsyl-OLIGOSACCHARI	6.35e+02	826	4	849	1	Y343_METJA	HYPOTHETICAL PROTEIN M	6.35e+02



827	4	25.0	1	RA54_SCHPO	DNA REPAIR PROTEIN RHP	6.35e+02	900	4	25.0	1064	1	YY08_METJA	HYPOTHETICAL PROTEIN M	6.35e+02
828	4	25.0	854	ENV_SIVAI	ENVELOPE POLYPROTEIN G	6.35e+02	901	4	25.0	1065	1	CERU_HUMAN	CERULOPLASMIN PRECURSOR	6.35e+02
829	4	25.0	857	V2A_CMVMB	PROBABLE RNA-DIRECTED	6.35e+02	902	4	25.0	1066	1	KL61_DROME	BIPOLAR KINESIN KRP-13	6.35e+02
830	4	25.0	857	V2A_CMVKB	PROBABLE RNA-DIRECTED	6.35e+02	903	4	25.0	1068	1	YCF0_MARPO	HYPOTHETICAL 127 KD PR	6.35e+02
831	4	25.0	857	V2A_CMVB	PROBABLE RNA-DIRECTED	6.35e+02	904	4	25.0	1103	1	DP0D_MESAU	DNA POLYMERASE DELTA C	6.35e+02
832	4	25.0	858	CIYAG_DICDI	ADENYLATE CYCLASE, GER	6.35e+02	905	4	25.0	1103	1	DP0D_RAT	DNA POLYMERASE DELTA C	6.35e+02
833	4	25.0	859	SULI_YEAST	SULFATE PERMEASE (HIGH	6.35e+02	906	4	25.0	1107	1	DP0D_HUMAN	DNA POLYMERASE DELTA C	6.35e+02
834	4	25.0	859	OSHI_YEAST	OSHI PROTEIN.	6.35e+02	907	4	25.0	1113	1	Y140_MYCPN	HYPOTHETICAL ATP-BINDI	6.35e+02
835	4	25.0	862	EAET_DICDI	ALPHA-ACTININ 3, NON M	6.35e+02	908	4	25.0	1122	1	Y140_MYCCE	HYPOTHETICAL 128.8 KD	6.35e+02
836	4	25.0	865	ENV_SIVAT	ENVELOPE POLYPROTEIN G	6.35e+02	909	4	25.0	1122	1	YG3C_YEAST	HYPOTHETICAL 128.8 KD	6.35e+02
837	4	25.0	866	MYSP_SCHMA	PARAMYOSIN.	6.35e+02	910	4	25.0	1135	1	NINS_DROME	NINAC SHORT PROTEIN (E	6.35e+02
838	4	25.0	869	YD95_YEAST	HYPOTHETICAL 98.7 KD P	6.35e+02	911	4	25.0	1145	1	PR22_YEAST	PRE-MRNA SPLICING FACT	6.35e+02
839	4	25.0	875	Y066_NVP0P	HYPOTHETICAL 98.6 KD P	6.35e+02	912	4	25.0	1146	1	MFD_HAEIN	TRANSCRIPTION-REPAIR C	6.35e+02
840	4	25.0	875	YDD4_SCHPO	HYPOTHETICAL 100.5 KD	6.35e+02	913	4	25.0	1153	1	YPDB_PLAKN	DUFFY RECEPTOR, BETA F	6.35e+02
841	4	25.0	877	ENV_SIVAG	ENVELOPE POLYPROTEIN G	6.35e+02	914	4	25.0	1159	1	YQ4_CAEEL	HYPOTHETICAL 127.4 KD	6.35e+02
842	4	25.0	878	YB9X_YEAST	HYPOTHETICAL 98.1 KD T	6.35e+02	915	4	25.0	1165	1	CHS3_YEAST	CHITIN SYNTHASE 3 (EC	6.35e+02
843	4	25.0	890	IF2_ECOLI	TRANSLATION INITIATION	6.35e+02	916	4	25.0	1172	1	CUI4_SCHPO	CHROMOSOME SEGREGATION	6.35e+02
844	4	25.0	894	YN86_YEAST	HYPOTHETICAL 99.5 KD P	6.35e+02	917	4	25.0	1173	1	VGL2_CVH22	E2 GLYCOPROTEIN PRECUR	6.35e+02
845	4	25.0	897	AACS_CHICK	ALPHA-ACTININ, SKELETA	6.35e+02	918	4	25.0	1178	1	TSF1_CHICK	THROMBOSPONDIN 1 PRECU	6.35e+02
846	4	25.0	903	DP0L_BPR69	DNA POLYMERASE (EC 2.7	6.35e+02	919	4	25.0	1186	1	XFG_HUMAN	DNA-REPAIR PROTEIN COM	6.35e+02
847	4	25.0	907	Y274_HUMAN	HYPOTHETICAL PROTEIN K	6.35e+02	920	4	25.0	1192	1	TOP2_ASF67	DNA TOPOISOMERASE II (	6.35e+02
848	4	25.0	915	PAC6_MOUSE	SERINE PROTEASE PC6 PR	6.35e+02	921	4	25.0	1200	1	DPOG_XENLA	DNA TOPOISOMERASE GAMMA (	6.35e+02
849	4	25.0	915	NFM_HUMAN	NEUROFILAMENT TRIPLET	6.35e+02	922	4	25.0	1201	1	YEO1_SCHPO	HYPOTHETICAL 138.5 KD	6.35e+02
850	4	25.0	925	VPH_BPRP1	PROBABLE TAIL FIBRE PR	6.35e+02	923	4	25.0	1203	1	SDC1_CAEEL	ZINC FINGER PROTEIN SD	6.35e+02
851	4	25.0	925	HEX_ADS41	HEXON PROTEIN (LATE PR	6.35e+02	924	4	25.0	1220	1	IF2P_HUMAN	PROBABLE TRANSLATION I	6.35e+02
852	4	25.0	933	SLAP_CAFME	S-LAYER PROTEIN (SUFRA	6.35e+02	925	4	25.0	1222	1	S160_YEAST	SCP160 PROTEIN (PROTEI	6.35e+02
853	4	25.0	935	YAC6_SCHPO	HYPOTHETICAL 105.6 KD	6.35e+02	926	4	25.0	1227	1	LAF4_HUMAN	LAF-4 PROTEIN (LYMPHOI	6.35e+02
854	4	25.0	937	PAC4_RAT	SUBTILISIN-LIKE PROTEA	6.35e+02	927	4						

973 4 25.0 1372 1 MYST\_RABIT 6.35e+02  
 974 4 25.0 1576 1 MYOIN HEAVY CHAIN, SM 6.35e+02  
 975 4 25.0 2014 1 YJUT\_YEAST 6.35e+02  
 976 4 25.0 2022 1 YJUT\_YEAST 6.35e+02  
 977 4 25.0 2051 1 FASL\_YEAST 6.35e+02  
 978 4 25.0 2052 1 YDOB\_SCHPO 6.35e+02  
 979 4 25.0 2133 1 FAS\_PIG 6.35e+02  
 980 4 25.0 2144 1 GLI1\_YEAST 6.35e+02  
 981 4 25.0 2214 1 PYL1\_YEAST 6.35e+02  
 982 4 25.0 2301 1 POLG\_TMEVD 6.35e+02  
 983 4 25.0 2318 1 NTC3\_MOUSE 6.35e+02  
 984 4 25.0 2483 1 PCX\_DROME 6.35e+02  
 985 4 25.0 2559 1 FAF\_MOUSE 6.35e+02  
 986 4 25.0 2670 1 YA05\_SCHPO 6.35e+02  
 987 4 25.0 2688 1 ZEP1\_MOUSE 6.35e+02  
 988 4 25.0 2747 1 FAF\_DROME 6.35e+02  
 989 4 25.0 2749 1 IP3R\_MOUSE 6.35e+02  
 990 4 25.0 3027 1 POLG\_PYFV1 6.35e+02  
 991 4 25.0 3144 1 HD\_HUMAN 6.35e+02  
 992 4 25.0 3148 1 HD\_FUGRU 6.35e+02  
 993 4 25.0 3175 1 RPOA\_EAV 6.35e+02  
 994 4 25.0 3432 1 POLG\_JAEV5 6.35e+02  
 995 4 25.0 3433 1 UTR0\_HUMAN 6.35e+02  
 996 4 25.0 3672 1 LML2\_CAEEL 6.35e+02  
 997 4 25.0 3770 1 ACVS\_Emeni 6.35e+02  
 998 4 25.0 3947 1 SID2\_TSTMA 6.35e+02  
 999 4 25.0 4385 1 YP73\_CAEEL 6.35e+02  
 1000 4 25.0 4544 1 LRPI\_HUMAN 6.35e+02

## ALIGNMENTS

RESULT 1  
 ID SCIP1\_BRALA STANDARD: PRT: 185 AA.  
 AC P04569;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE SARCOPLASMIC CALCIUM-BINDING PROTEINS I, III, AND IV (SCP I, III, IV).  
 OS BRANCHIOSTOMA LANCEOLATUM (COMMON LANCELET) (AMPHIOXUS).  
 OC EUKARYOTA; METAZOA; CHORDATA; CEPHALOCHORDATA; BRANCHIOSTOMIDAE;  
 OC BRANCHIOSTOMA.  
 RN [1]  
 RP SEQUENCE (SCP I).  
 RA TAKAGI T., KONISHI K., COX J.A.;  
 RT "Amino acid sequence of two sarcoplasmic calcium-binding proteins  
 from the protochordate amphioxus.";  
 RL BIOCHEMISTRY 25:3585-3592(1986).  
 RN [2]  
 RP SEQUENCE (SCP III AND IV), AND REVISIONS (SCP I).  
 RX MEDLINE: 91006125.  
 RA TAKAGI T., COX J.A.;  
 RT "Amino acid sequences of four isoforms of amphioxus sarcoplasmic  
 calcium-binding proteins.";  
 RL EUR. J. BIOL. 192:387-399(1990).  
 CC -!- FUNCTION: LIKE PARVALBUMINS, SCP'S SEEM TO BE MORE ABUNDANT IN  
 FAST CONTRACTING MUSCLES, BUT NO FUNCTIONAL RELATIONSHIP CAN BE  
 ESTABLISHED FROM THIS DISTRIBUTION.  
 CC -!- THE SARCOPLASMIC CALCIUM-BINDING PROTEINS ARE ABUNDANT IN THE  
 MUSCLE OF ARTHROPODS, MOLLUSCS, ANNELIDS, AND PROTOCHORDATES.  
 CC -!- THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING SITES;  
 POTENTIAL SITE 4 HAS LOST AFFINITY FOR CALCIUM.  
 CC -!- THERE ARE 7 DIFFERENT SCP'S IN AMPHIOXUS. THE SEQUENCE SHOWN  
 HERE IS THAT OF SCP I.  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
 DR PIR: B24479; B24479.  
 DR PIR: S13182; S13182.  
 DR PIR: S13184; S13184.  
 DR PIR: S13185; S13185.  
 DR PROSITE: PS00018; EF\_HAND; 3.  
 DR PFAM: PF00036; ehand; 1.  
 DR HSSP: P04570; 2SAS.  
 KW MUSCLE PROTEIN; CALCIUM-BINDING; DUPLICATION.

FT CA\_BIND 19 30 SITE 1.  
 FT CA\_BIND 70 81 SITE 2.  
 FT CA\_BIND 115 126 SITE 3.  
 FT DOMAIN 150 161 ANCESTRAL CALCIUM SITE 4.  
 FT VARIANT 20 20 Y -> M (IN SCP III).  
 FT VARIANT 23 23 D -> N (IN SCP IV).  
 SQ SEQUENCE 185 AA; 21418 MW; B0BC7795 CRC32;  
 Query Match 37.5%; Score 6; DB 1; Length 185;  
 Best Local Similarity 100.0%; Pred.No. 2.43e-01; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0;  
 Db 71 INKDDV 76  
 QY 2 INKDDV 7  
 RESULT 2  
 ID SCIP2\_BRALA STANDARD: PRT: 185 AA.  
 AC P04570;  
 DT 13-AUG-1987 (REL. 05, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
 DE SARCOPLASMIC CALCIUM-BINDING PROTEINS II, V, VI, AND VII (SCP II, V, VI, VII).  
 OS BRANCHIOSTOMA LANCEOLATUM (COMMON LANCELET) (AMPHIOXUS).  
 OC EUKARYOTA; METAZOA; CHORDATA; CEPHALOCHORDATA; BRANCHIOSTOMIDAE;  
 OC BRANCHIOSTOMA.  
 RN [1]  
 RP SEQUENCE.  
 RA TAKAGI T., KONISHI K., COX J.A.;  
 RT "Amino acid sequence of two sarcoplasmic calcium-binding proteins  
 from the protochordate amphioxus.";  
 RL BIOCHEMISTRY 25:3585-3592(1986).  
 RN [2]  
 RP REVISIONS.  
 RX MEDLINE: 91006125.  
 RA TAKAGI T., COX J.A.;  
 RT "Amino acid sequences of four isoforms of amphioxus sarcoplasmic  
 calcium-binding proteins.";  
 RL EUR. J. BIOL. 192:387-399(1990).  
 RN [3]  
 RP SEQUENCE (SCP V TO VII).  
 RX MEDLINE: 92339504.  
 RA TAKAGI T., VALETTE-TALBI L., COX J.A.;  
 RT "Primary structure of three minor isoforms of amphioxus sarcoplasmic  
 calcium-binding proteins.";  
 RL FEBS LETT. 302:159-160(1992).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE: 93156054.  
 RA COOK W.J., JEFFREY L.C., COX J.A., VIJAY-KUMAR S.;  
 RT "Structure of a sarcoplasmic calcium-binding protein from amphioxus  
 refined at 2.4-A resolution.";  
 RL J. MOL. BIOL. 229:461-471(1993).  
 CC -!- FUNCTION: LIKE PARVALBUMINS, SCP'S SEEM TO BE MORE ABUNDANT IN  
 FAST CONTRACTING MUSCLES, BUT NO FUNCTIONAL RELATIONSHIP CAN BE  
 ESTABLISHED FROM THIS DISTRIBUTION.  
 CC -!- THE SARCOPLASMIC CALCIUM-BINDING PROTEINS ARE ABUNDANT IN THE  
 MUSCLE OF ARTHROPODS, MOLLUSCS, ANNELIDS, AND PROTOCHORDATES.  
 CC -!- THIS PROTEIN HAS THREE FUNCTIONAL CALCIUM-BINDING SITES;  
 POTENTIAL SITE 4 HAS LOST AFFINITY FOR CALCIUM.  
 CC -!- THERE ARE 7 DIFFERENT SCP'S IN AMPHIOXUS. THE SEQUENCE SHOWN  
 HERE IS THAT OF SCP II.  
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.  
 DR PIR: A24479; A24479.  
 DR PIR: S13183; S13183.  
 DR PDB: 2SAS; 31-OCT-93.  
 DR PROSITE: PS00018; EF\_HAND; 3.  
 DR PFAM: PF00036; ehand; 1.  
 KW MUSCLE PROTEIN; CALCIUM-BINDING; DUPLICATION; 3D-STRUCTURE.  
 FT CA\_BIND 19 30 SITE 1.  
 FT CA\_BIND 70 81 SITE 2.

FT CA\_BIND 115 126 SITE 3.  
FT DOMAIN 150 161 ANCESTRAL CALCIUM SITE 4.  
FT VARIANT 25 28 S -> A (IN SCP V).  
FT VARIANT 28 25 D -> W (IN SCP V, VI, AND VII).  
FT VARIANT 29 29 N -> E (IN SCP VI).  
FT HELIX 4 17  
FT TURN 18 18  
FT TURN 20 21  
FT STRAND 25 26  
FT HELIX 28 45  
FT HELIX 48 69  
FT TURN 71 72  
FT STRAND 77 78  
FT HELIX 79 91  
FT TURN 92 92  
FT TURN 92 98  
FT TURN 101 102  
FT HELIX 103 114  
FT TURN 116 117  
FT HELIX 124 130  
FT TURN 131 132  
FT HELIX 141 149  
FT TURN 150 153  
FT HELIX 158 170  
FT TURN 176 181  
FT TURN 182 182  
SQ SEQUENCE 185 AA; 21286 MW; 20716924 CRC32;  
  
Query Match 37.5%; Score 6; DB 1; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.43e-01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 71 INKDDV 76  
QY 2 INKDDV 7  
  
RESULT 3  
ID SR72\_CANFA STANDARD: PRT: 670 AA.  
AC P33731;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE SIGNAL RECOGNITION PARTICLE 72 KD PROTEIN (SRP72).  
GN SRP72.  
OS CANIS FAMILIARIS (DOG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC CARNIVORA; FISSIPEDIA; CANIDAE; CANIS.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE; 93273803.  
RA LUETCKE H., PREHN S., ASHFORD A.J., REMUS M., FRANK R.,  
RA DOBERSTEIN B.;  
RT "Assembly of the 68- and 72-kD proteins of signal recognition particle with 7S RNA.";  
RL J. CELL BIOL. 121:977-985(1993).  
CC -!- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY HAS A CRUCIAL ROLE IN TARGETING SECRETORY PROTEINS TO THE ROUGH ENDOPLASMIC RETICULUM MEMBRANE. SRP72 BINDS THE 7S RNA ONLY IN PRESENCE OF SRP68. THIS RIBONUCLEOPROTEIN COMPLEX MIGHT INTERACT DIRECTLY WITH THE DOCKING PROTEIN IN THE ER MEMBRANE AND POSSIBLY PARTICIPATE IN THE ELONGATION ARREST FUNCTION.  
CC -!- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA MOLECULE SRP19, SRP14 AND SRP9.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- DOMAIN: THE C-TERMINAL IS ESSENTIAL FOR THE INTERACTION WITH THE SRP68/7S RNA COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE SRP72 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X67813; G297768; -.  
DR PIR; S32167; S32167.  
DR PIR; A40692; A40692.  
KW SIGNAL RECOGNITION PARTICLE; RIBONUCLEOPROTEIN.  
FT INIT\_MET 0 PROBABLE.  
FT MOD\_RES 1 560 BLOCKED.  
FT DOMAIN 551 560 POLY-LYS.  
FT DOMAIN 661 664 POLY-LYS.  
SQ SEQUENCE 670 AA; 74362 MW; 9CB38962 CRC32;  
  
Query Match 37.5%; Score 6; DB 1; Length 670;  
Best Local Similarity 100.0%; Pred. No. 2.43e-01;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 39 INKDDV 44  
QY 2 INKDDV 7  
  
RESULT 4  
ID SR72\_HUMAN STANDARD: PRT: 670 AA.  
AC O76094;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE SIGNAL RECOGNITION PARTICLE 72 KD PROTEIN (SRP72).  
GN SRP72.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA GOWDA K., ZWIEB C.;  
RT "Protein SRP72 sequence of human signal recognition particle.";  
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA UTZ P.J., HOTELET M., MILLER I.J., ANDERSON P.;  
RT "Sequence of human signal recognition particle (SRP) 72.";  
RL SUBMITTED (JUN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: SIGNAL-RECOGNITION-PARTICLE ASSEMBLY HAS A CRUCIAL ROLE IN TARGETING SECRETORY PROTEINS TO THE ROUGH ENDOPLASMIC RETICULUM MEMBRANE. SRP72 BINDS THE 7S RNA ONLY IN PRESENCE OF SRP68. THIS RIBONUCLEOPROTEIN COMPLEX MIGHT INTERACT DIRECTLY WITH THE DOCKING PROTEIN IN THE ER MEMBRANE AND POSSIBLY PARTICIPATE IN THE ELONGATION ARREST FUNCTION.  
CC -!- SUBUNIT: SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA MOLECULE OF 300 NUCLEOTIDES AND SIX PROTEIN SUBUNITS: SRP72, SRP68, SRP54, SRP19, SRP14 AND SRP9.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- DOMAIN: THE C-TERMINAL IS ESSENTIAL FOR THE INTERACTION WITH THE SRP68/7S RNA COMPLEX (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SRP72 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AF077019; G3335650; -.  
DR EMBL; AF069765; G3243033; -.  
DR MIM; 602122; -.  
KW SIGNAL RECOGNITION PARTICLE; RIBONUCLEOPROTEIN.  
FT INIT\_MET 0  
FT DOMAIN 551 560 POLY-LYS.  
CC

```
FT DOMAIN 661 664 POLY-LYS.
SQ SEQUENCE 670 AA; 74475 MW; 2EFP8AB8 CRC32;

Query Match 37.5%; Score 6; DB 1; Length 670;
Best Local Similarity 100.0%; Pred. No. 2.43e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 INKDDV 44
|||||
Qy 2 INKDDV 7

RESULT 5
ID MGR3_HUMAN STANDARD; PRT; 877 AA.
AC Q14832;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE METABOTROPIC GLUTAMATE RECEPTOR 3 PRECURSOR.
GN GRM3 OR MGLUR3.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 96437205.
RA MAKOFF A., VOLPE F., LELCHUK R., HARRINGTON K., EMSON P.;
RT "Molecular characterization and localization of human metabotropic
RL glutamate receptor type 3.";
RL BRAIN RES. MOL. BRAIN RES. 40:55-63(1996).
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- STRONGEST, TO MGLUR2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X77748; E99039; -
DR GCRDB; GCR_2070; -
DR MIM; 601115; -
DR PROSITE; PS00979; G_PROTEIN_RECF_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECF_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECF_F3_3; 1.
DR PFAM; PF00003; 7tm_3; 1.
DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL;
KW MULTIGENE FAMILY.
FT SIGNAL 1 20
FT CHAIN 21 877
FT DOMAIN 21 574
FT TRANSMEM 575 597
FT DOMAIN 598 611
FT TRANSMEM 612 632
FT DOMAIN 633 643
FT TRANSMEM 644 662
FT DOMAIN 663 686
FT TRANSMEM 687 707
FT DOMAIN 708 732
FT TRANSMEM 733 754
FT DOMAIN 755 767
FT TRANSMEM 768 790
FT DOMAIN 791 800
FT TRANSMEM 801 826
FT DOMAIN 827 877
FT CARBOHYD 207

Query Match 37.5%; Score 6; DB 1; Length 670;
Best Local Similarity 100.0%; Pred. No. 2.43e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 INKDDV 44
|||||
Qy 2 INKDDV 7

RESULT 6
ID CLH_CABEL STANDARD; PRT; 1681 AA.
AC P34574;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE CLATHRIN HEAVY CHAIN.
GN T20G5.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA BERKS M., SMITH A.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: CLATHRIN IS THE MAJOR PROTEIN OF THE POLYHEDRAL COAT OF
CC COATED PITS AND VESICLES (BY SIMILARITY).
CC -!- SUBUNIT: CLATHRIN TRISKELIONS, COMPOSED OF 3 HEAVY CHAINS AND 3
CC LIGHT CHAINS, ARE THE BASIC SUBUNITS OF THE CLATHRIN COAT
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC FACE OF COATED PITS AND
CC VESICLES (BY SIMILARITY).
CC -!- DOMAIN: THE C-TERMINAL THIRD OF THE HEAVY CHAINS FORMS THE HUB OF
CC THE TRISKELION. THIS REGION CONTAINS THE TRIMERIZATION DOMAIN AND
CC THE LIGHT-CHAIN BINDING DOMAIN INVOLVED IN THE ASSEMBLY OF THE
CC CLATHRIN LATTICE.
CC -!- SIMILARITY: BELONGS TO THE CLATHRIN HEAVY CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z30423; G458481; -
DR PIR; S42369; S42369.
DR WORMPEP; T20G5.1; CE00480.
DR PFAM; PF00637; Clathrin_repeat; 7.
DR KW HYPOTHETICAL PROTEIN; COATED PITS.
SQ SEQUENCE 1681 AA; 191541 MW; 6CD2B462 CRC32;

Query Match 37.5%; Score 6; DB 1; Length 1681;
Best Local Similarity 100.0%; Pred. No. 2.43e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1256 KQVCFS 1261
|||||
Qy 10 KQVCFS 15

RESULT 7
ID RPPO_PVMR STANDARD; PRT; 1967 AA.
AC P17965;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
```



```

488  DOMAIN 524  EGF-LIKE 12,  CALCIIUM-BINDING (POTENTIAL).
489  DOMAIN 526  EGF-LIKE 13,  CALCIIUM-BINDING (POTENTIAL).
490  DOMAIN 564  EGF-LIKE 14,  CALCIIUM-BINDING (POTENTIAL).
491  DOMAIN 600  EGF-LIKE 15,  CALCIIUM-BINDING (POTENTIAL).
492  DOMAIN 602  EGF-LIKE 16,  CALCIIUM-BINDING (POTENTIAL).
493  DOMAIN 637  EGF-LIKE 17,  CALCIIUM-BINDING (POTENTIAL).
494  DOMAIN 677  EGF-LIKE 18,  CALCIIUM-BINDING (POTENTIAL).
495  DOMAIN 715  EGF-LIKE 19,  CALCIIUM-BINDING (POTENTIAL).
496  DOMAIN 751  EGF-LIKE 20,  CALCIIUM-BINDING (POTENTIAL).
497  DOMAIN 789  EGF-LIKE 21,  CALCIIUM-BINDING (POTENTIAL).
498  DOMAIN 829  EGF-LIKE 22,  CALCIIUM-BINDING (POTENTIAL).
499  DOMAIN 865  EGF-LIKE 23,  CALCIIUM-BINDING (POTENTIAL).
500  DOMAIN 905  EGF-LIKE 24,  CALCIIUM-BINDING (POTENTIAL).
501  DOMAIN 946  EGF-LIKE 25,  CALCIIUM-BINDING (POTENTIAL).
502  DOMAIN 984  EGF-LIKE 26,  CALCIIUM-BINDING (POTENTIAL).
503  DOMAIN 1022  EGF-LIKE 27,  CALCIIUM-BINDING (POTENTIAL).
504  DOMAIN 1060  EGF-LIKE 28,  CALCIIUM-BINDING (POTENTIAL).
505  DOMAIN 1098  EGF-LIKE 29,  CALCIIUM-BINDING (POTENTIAL).
506  DOMAIN 1136  EGF-LIKE 30,  CALCIIUM-BINDING (POTENTIAL).
507  DOMAIN 1183  EGF-LIKE 31,  CALCIIUM-BINDING (POTENTIAL).
508  DOMAIN 1221  EGF-LIKE 32,  CALCIIUM-BINDING (POTENTIAL).
509  DOMAIN 1259  EGF-LIKE 33,  CALCIIUM-BINDING (POTENTIAL).
510  DOMAIN 1297  EGF-LIKE 34,  CALCIIUM-BINDING (POTENTIAL).
511  DOMAIN 1337  EGF-LIKE 35,  CALCIIUM-BINDING (POTENTIAL).
512  DOMAIN 1375  EGF-LIKE 36,  CALCIIUM-BINDING (POTENTIAL).
513  DOMAIN 1415  EGF-LIKE 37,  CALCIIUM-BINDING (POTENTIAL).
514  DOMAIN 1451  EGF-LIKE 38,  CALCIIUM-BINDING (POTENTIAL).
515  DOMAIN 1475  EGF-LIKE 39,  CALCIIUM-BINDING (POTENTIAL).
516  DOMAIN 1513  EGF-LIKE 40,  CALCIIUM-BINDING (POTENTIAL).
517  REPEAT 1475  LIN/NOTCH 1.
518  REPEAT 1514  LIN/NOTCH 2.
519  REPEAT 1554  LIN/NOTCH 3.
520  DOMAIN 1896  POLY-GLN (OPA-REPEAT).
521  DOMAIN 2538  POLY-GLN (OPA-REPEAT).
522  DOMAIN 62  73  BY SIMILARITY.
523  DOMAIN 67  83  BY SIMILARITY.
524  DOMAIN 85  94  BY SIMILARITY.
525  DOMAIN 100  111  BY SIMILARITY.
526  DOMAIN 105  124  BY SIMILARITY.
527  DOMAIN 126  135  BY SIMILARITY.
528  DOMAIN 143  154  BY SIMILARITY.
529  DOMAIN 148  164  BY SIMILARITY.
530  DOMAIN 166  175  BY SIMILARITY.
531  DOMAIN 181  192  BY SIMILARITY.
532  DOMAIN 203  214  BY SIMILARITY.
533  DOMAIN 205  214  BY SIMILARITY.
534  DOMAIN 221  232  BY SIMILARITY.
535  DOMAIN 226  241  BY SIMILARITY.
536  DOMAIN 243  252  BY SIMILARITY.
537  DOMAIN 259  270  BY SIMILARITY.
538  DOMAIN 264  279  BY SIMILARITY.
539  DOMAIN 281  290  BY SIMILARITY.
540  DOMAIN 297  308  BY SIMILARITY.
541  DOMAIN 302  317  BY SIMILARITY.
542  DOMAIN 319  328  BY SIMILARITY.
543  DOMAIN 335  349  BY SIMILARITY.
544  DOMAIN 343  358  BY SIMILARITY.
545  DOMAIN 360  369  BY SIMILARITY.
546  DOMAIN 376  387  BY SIMILARITY.
547  DOMAIN 381  396  BY SIMILARITY.
548  DOMAIN 398  407  BY SIMILARITY.
549  DOMAIN 413  424  BY SIMILARITY.
550  DOMAIN 418  435  BY SIMILARITY.
551  DOMAIN 437  446  BY SIMILARITY.
552  DOMAIN 453  465  BY SIMILARITY.
553  DOMAIN 459  474  BY SIMILARITY.
554  DOMAIN 476  485  BY SIMILARITY.
555  DOMAIN 492  503  BY SIMILARITY.
556  DOMAIN 497  512  BY SIMILARITY.
557  DOMAIN 514  523  BY SIMILARITY.
558  DOMAIN 530  541  BY SIMILARITY.
559  DOMAIN 535  550  BY SIMILARITY.
560  DOMAIN 552  561  BY SIMILARITY.
561  DOMAIN 568  579  BY SIMILARITY.
562  DOMAIN 573  588  BY SIMILARITY.
563  DOMAIN 590  599  BY SIMILARITY.

...
Note: remainder of annotations omitted.

Query Match 37.5%; Score 6; DB 1; Length 2703;
Best Local Similarity 100.0%; Pred. No. 2.43e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1257 EINKDD 1262
QY 1 EINKDD 6

RESULT 9
ID Y02B-LPT4 STANDARD; PRT; 60 AA.
AC P39232;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 7.0 KD PROTEIN IN GP58-GP41 INTERGENIC REGION.
GN Y02B OR 61.5.
OS BACTERIOPHAGE T4.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE;
OC T4-LIKE PHAGES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93188183.
RA SELICK H.E., STORMO G.D., DYSON R.L., ALBERTS B.M.;
RT "Analysis of five presumptive protein-coding sequences clustered
between the primosome genes, 41 and 61, of bacteriophages T4, T2, and
T6."
RL J. VIROL. 67:2305-2316(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S57514; G298519; -
CC PIR; A45681; A45681.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 60 AA; 7027 MW; F71C4EAC CRC32;

Query Match 31.3%; Score 5; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 26 KDDVQ 30
QY 4 KDDVQ 8
```

```

RX MEDLINE: 96408771.
RA KATSOULOU C., TZERNIA M., TAVERNARAKIS N., ALEXANDRAKI D.:
RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT chromosome X reveals 14 known genes and 13 new open reading frames
RT including homologues of genes clustered on the right arm of
RT chromosome XI."
RL YEAST 12:787-797(1996).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE UPF0057 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J04453; G806316; -
CC EMBL: Z49427; G1015369; -
CC DR EMBL: X87371; G854546; -
CC DR PIR: A32209; A32209.
CC DR PIR: B30902; B30902.
CC DR PROSITE: PS01309; UPF0057. 1.
CC KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
CC FT TRANSMEM 17 37 POTENTIAL.
CC FT TRANSMEM 49 69 POTENTIAL.
CC SQ SEQUENCE 133 AA; 15199 MW; BELL1BB4 CRC32;

Query Match 31.3%; Score 5; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 INKDD 21
   |||||
QY 2 INKDD 6

RESULT 12
ID Y391_METJA STANDARD; PRT; 183 AA.
AC Q57836;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0391.
GN MJ0391.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
CC METHANOCOCCUS.
CC [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.:
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL JANNASCHII.
CC -!- SIMILARITY: SOME, TO S.TYPHIMURIUM PRECORRIN-8W DECARBOXYLASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67496; G1591157; -
CC TIGR: MJ0453; -
CC KW HYPOTHETICAL PROTEIN.
CC SQ SEQUENCE 107 AA; 12603 MW; 05EDFF44 CRC32;

Query Match 31.3%; Score 5; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 EINKD 85
   |||||
QY 1 EINKD 5

RESULT 11
ID YJPI_YEAST STANDARD; PRT; 133 AA.
AC P14359;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 15.2 KD PROTEIN IN INO1-IDS2 INTERGENIC REGION.
GN YJL151C OR J0630
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
CC SACCCHAROMYCETACEAE; SACCCHAROMYCETES.
CC [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89093118.
RA DEAN-JOHNSON M., HENRY S.A.:
RT "Biosynthesis of inositol in yeast. Primary structure of
RT myo-inositol-1-phosphate synthase (EC 5.5.1.4) and functional
RT analysis of its structural gene, the INO1 locus."
RL J. BIOL. CHEM. 264:1274-1283(1989).
CC [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;

```

```
-----
CC EMBL; U67491; G1591095; -.
DR TIGR; MJ0391; -.
DR PFAM; PF00398; RnaAD; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 183 AA; 20502 MW; C6ABE676 CRC32;

Query Match          31.3%; Score 5; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 NKDDV 38
|||||
QY 3 NKDDV 7

RESULT 13
ID TPIS_METJA STANDARD; PRT; 219 AA.
AC Q58923;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PUTATIVE TRIOSEPHOSPHATE ISOMERASE (EC 5.3.1.1) (TIM).
GN TPIA OR MJ1528.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURVARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
CC [1]
CC SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RC MEDLINE; 96337999.
RX BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEORGHAN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL SCIENCE 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE = DIHYDROXY-
CC ACETONE PHOSPHATE.
CC -!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67593; G1592158; -.
DR TIGR; MJ1528; -.
DR PROSITE; PS00171; TIM; FALSE_NEG.
DR HSP; P00929; 2TSY.
KW ISOMERASE; GLYCOLYSIS; GLUCONEOGENESIS; FATTY ACID BIOSYNTHESIS;
KW PENTOSE SHUNT.
FT ACT_SITE 90 90 BY SIMILARITY.
SQ SEQUENCE 219 AA; 23319 MW; A0A3BDEC CRC32;

Query Match          31.3%; Score 5; DB 1; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 166 EINKD 170
|||||
QY 1 EINKD 5

-----
```

```
-----
RESULT 14
ID P1MT_BRARE STANDARD; PRT; 227 AA.
AC Q92047;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROTEIN-L-ISOASPARTATE(D-ASPARTATE) O-METHYLTRANSFERASE (EC 2.1.1.77)
DE (PROTEIN-BETA-ASPARTATE METHYLTRANSFERASE) (P1MT) (PROTEIN L-
DE ISOASPARTYL/D-ASPARTYL METHYLTRANSFERASE) (L-ISOASPARTYL PROTEIN
DE CARBOXYL METHYLTRANSFERASE).
GN PCMT.
OS BRACHYDANIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
OC TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;
OC CYPRINIDAE; RASBORINAE; DANIO.
CC [1]
CC SEQUENCE FROM N.A.
RP MEDLINE; 97397023.
RX KAGAN R.M., MCFADDEN H.J., MCFADDEN P.N., O'CONNOR C., CLARKE S.;
RT "Molecular phylogenetics of a protein repair methyltransferase.";
RL COMP. BIOCHEM. PHYSIOL. 117B:379-395(1997).
CC -!- FUNCTION: CATALYZES THE METHYL ESTERIFICATION OF L-ISOASPARTYL AND
CC D-ASPARTYL RESIDUES IN PEPTIDES AND PROTEINS THAT RESULT FROM
CC SPONTANEOUS DECOMPOSITION OF NORMAL L-ASPARTYL AND L-ASPARAGINYL
CC RESIDUES. IT PLAYS A ROLE IN THE REPAIR AND/OR DEGRADATION OF
CC DAMAGED PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + PROTEIN L-BETA-
CC ASPARTATE = S-ADENOSYL-L-HOMOCYSTEINE + PROTEIN L-BETA-ASPARTATE
CC METHYL ESTER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L-ISOASPARTYL/D-ASPARTYL PROTEIN
CC METHYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U37434; G1255009; -.
DR PROSITE; PS01279; PCMT; 1.
DR PFAM; PF01135; PCMT; 1.
KW TRANSFERASE; METHYLTRANSFERASE; ACETYLATION.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 227 AA; 24594 MW; B8C76FD6 CRC32;

Query Match          31.3%; Score 5; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 93 VCFSR 97
|||||
QY 12 VCFSR 16

RESULT 15
ID SEM5_CAEEL STANDARD; PRT; 228 AA.
AC P29355;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SEX MUSCLE ABNORMAL PROTEIN 5.
GN SEM-5 OR C14F5.5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
CC [1]
CC SEQUENCE FROM N.A.
RP MEDLINE; 92195405.
RX CLARK S.G., STERN M.J., HORVITZ H.R.;
```



```
RT "C. elegans cell-signalling gene sem-5 encodes a protein with SH2 and
RT SH3 domains.";
RL NATURE 356:340-344(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP STRUCTURE BY NMR OF SECOND SH3 DOMAIN.
RX MEDLINE; 95059481.
RA LIM W.A., RICHARDS F.M., FOX R.O.;
RT "Structural determinants of peptide-binding orientation and of
RT sequence specificity in SH3 domains.";
RL NATURE 372:375-379(1994).
CC -!- FUNCTION: ACTS BOTH IN VULVAL INDUCTION AND SEX MYOBLAST
CC MIGRATION. SEM-5 PRESUMABLY INTERACTS WITH THE KINASE RECEPTOR
CC LET-23 AND WITH A TARGET THAT MODIFIES THE RAS-LIKE PROTEIN
CC LET-60.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; S88446; G247605; -.
DR EMBL; U29082; G861389; -.
DR PIR; S25730; S25730.
DR PDB; 1SEM; 10-JUL-95.
DR WORMPEP; C14F5.5; CE01784.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 2.
DR PFAM; PF00017; SH2; 1.
DR PFAM; PF00018; SH3; 2.
KW SH2 DOMAIN; SH3 DOMAIN; 3D-STRUCTURE.
FT DOMAIN 1 58
FT DOMAIN 60 152
FT DOMAIN 154 213
FT VARIANT 49 49 P -> L (IN SEM-5 MUTANT N1619).
FT VARIANT 90 90 E -> K (IN SEM-5 MUTANT N1779).
FT VARIANT 91 91 S -> N (IN SEM-5 MUTANT N1781).
FT VARIANT 201 201 G -> R (IN SEM-5 MUTANT N2195).
SQ SEQUENCE 228 AA; 26210 MW; 501A1C1C CRC32;

Query Match 31.3%; Score 5; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.64e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 184 INKDD 188
QY 2 INKDD 6
```

Search completed: Sat Aug 28 15:10:59 1999  
Job time : 46 secs.

**This Page Blank (uspto)**

\*\*\*\*\*  
WISREH  
\*\*\*\*\*  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution Rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 15:11:19 1999; MasPar time 7.29 Seconds  
119.826 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-49  
Description: (1-16) from US09049696.pgp  
Perfect Score: 16  
Sequence: 1 EINKDDVQHKQVCFSR 16

Scoring table:  
TABLE unitprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: sptrembl9  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phase 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 2.307; Variance 0.334; scale 6.912

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Match %	Description	ID	
1	6	37.5	440A LONG HYPOTHETICA	058162	3.59e-01
2	6	37.5	HYPOTHETICAL PROTEIN M	058327	3.59e-01
3	6	37.5	ORF1 223K PROTEIN.	089548	3.59e-01
4	6	37.5	NOTCH HOMOLOG.	016004	3.59e-01
5	5	31.3	ASP13 MADS BOX (FRAGME	038747	2.62e+01
6	5	31.3	ASP8 MADS BOX (FRAGMEN	038753	2.62e+01
7	5	31.3	REPLICATION-ASSOCIATED	067603	2.62e+01
8	5	31.3	TAT PROTEIN.	092083	2.62e+01
9	5	31.3	BETA-LACTAMASE (EC 3.5	053698	2.62e+01
10	5	31.3	CONSERVED HYPOTHETICAL	025775	2.62e+01
11	5	31.3	ORF141 PROTEIN.	024273	2.62e+01
12	5	31.3	HYPOTHETICAL 17.2 KD P	005984	2.62e+01
13	5	31.3	FOLYGLUTAMATE SYN	025308	2.62e+01
14	5	31.3	MYOSIN REGULATORY LIGH	030452	2.62e+01
15	5	31.3	HYPOTHETICAL 19.4 KD P	038750	2.62e+01
16	5	31.3	GAM.	087650	2.62e+01
17	5	31.3	T20K24.20 PROTEIN.	064482	2.62e+01
18	5	31.3	FROM BASES 2468374 TO	076514	2.62e+01
19	5	31.3	CHLORIN REDUCTASE SUBU	030837	2.62e+01
20	5	31.3	MC0989.	098265	2.62e+01

DJ370M22.1 (GROWTH FAC	2.62e+01
MEMBRAN-BOUND SMALL G	2.62e+01
REPLICATIVE PROTEIN (F	2.62e+01
REPLICATIVE PROTEIN (F	2.62e+01
U22660.	2.62e+01
GUANINE NUCLEOTIDE-BIN	2.62e+01
TYPE B BETA-LACTAMASE	2.62e+01
MOCC (FRAGMENT).	2.62e+01
PLASMID DNA FOR VANCOM	2.62e+01
ENDO-1,3-BETA-GLUCANAS	2.62e+01
COMPLETE GENOME.	2.62e+01
CHROMOSOME ARGININE TR	2.62e+01
BETA-LACTAMASE (EC 3.5	2.62e+01
M153.1 PROTEIN.	2.62e+01
BETA-LACTAMASE (EC 3.5	2.62e+01
BETA-LACTAMASE (EC 3.5	2.62e+01
CONSERVED HYPOTHETICAL	2.62e+01
34 KDA MOV34 ISOLOGUE.	2.62e+01
PLASMID PPL1, COMPLETE	2.62e+01
GENOME, PARTIAL SEQUEN	2.62e+01
W10C8.1 PROTEIN.	2.62e+01
SIMILARITY TO MOUSE MI	2.62e+01
K01D12.1 PROTEIN.	2.62e+01
322AA LONG HYPOTHETICA	2.62e+01
GRB-2-RELATED MONOCYTI	2.62e+01
COP9 COMPLEX SUBUNIT 6	2.62e+01
HYPOTHETICAL 37.3 KD P	2.62e+01
F24L7.15 PROTEIN.	2.62e+01
GADS PROTEIN.	2.62e+01
HYPOTHETICAL PROTEIN M	2.62e+01
T07G12.12 PROTEIN.	2.62e+01
STRAIN TAMAULIPAS COMP	2.62e+01
F41D3.3 PROTEIN.	2.62e+01
REPLICASE ASSOCIATED P	2.62e+01
CYTOSINE METHYLTRANSFER	2.62e+01
ERPM.	2.62e+01
SAGA.	2.62e+01
ALCOHOL DEHYDROGENASE.	2.62e+01
PROTEIN PHOSPHATASE 2C	2.62e+01
R09D1.13 PROTEIN.	2.62e+01
ORF393 PROTEIN.	2.62e+01
396AA LONG HYPOTHETICA	2.62e+01
MG2+ DEPENDENT PROTEIN	2.62e+01
PROTEIN PHOSPHATASE-2C	2.62e+01
HYPOTHETICAL PROTEIN M	2.62e+01
HAT A1.	2.62e+01
CARBOXY-TERMINAL PROCE	2.62e+01
N-ACETYLURAMOYL-L-ALA	2.62e+01
ORF 3.5 (FRAGMENT).	2.62e+01
F6E13.27 PROTEIN.	2.62e+01
ASPARTATE AMINOTRANSFE	2.62e+01
PP2C PROTEIN.	2.62e+01
F18A8.14 PROTEIN.	2.62e+01
K05F6.7 PROTEIN.	2.62e+01
CODED FOR BY C. ELEGAN	2.62e+01
SECRETION PROTEIN.	2.62e+01
HYPOTHETICAL 67.1 KD P	2.62e+01
RYANODINE RECEPTOR TYP	2.62e+01
HEAT SHOCK PROTEIN 70	2.62e+01
HYPOTHETICAL 71.8 KD P	2.62e+01
HSP70.	2.62e+01
TRANSMEMBRANE PROTEIN.	2.62e+01
F37C4.2 PROTEIN.	2.62e+01
C47E8.7 PROTEIN.	2.62e+01
ARABINOSIDASE.	2.62e+01
D8035.28P.	2.62e+01
ENVELOPE POLYPROTEIN.	2.62e+01
PYRUVATE, ORTHOPHOSPHAT	2.62e+01
F28C1.3 PROTEIN.	2.62e+01
PYRUVATE, ORTHOPHOSPHAT	2.62e+01
F17B5.2 PROTEIN.	2.62e+01
PUTATIVE CU2+-TRANSPOR	2.62e+01
ORNITHINE DECARBOXYLAS	2.62e+01

94	5	31.3	948	2	O30483	PUTATIVE REGULATORY PR	2.62e+01	167	4	25.0	72	14	O89726	TAT PROTEIN (FRAGMENT)	1.12e+03
95	5	31.3	962	14	O89443	HELICASE.	2.62e+01	168	4	25.0	72	14	O89720	TAT PROTEIN (FRAGMENT)	1.12e+03
96	5	31.3	974	13	O91735	EPH RECEPTOR TYROSINE	2.62e+01	169	4	25.0	72	14	O89719	TAT PROTEIN (FRAGMENT)	1.12e+03
97	5	31.3	1008	5	O77391	MAL3P6.4 PROTEIN.	2.62e+01	170	4	25.0	72	14	O89723	TAT PROTEIN (FRAGMENT)	1.12e+03
98	5	31.3	1098	2	O48152	ADHESIN (HIA).	2.62e+01	171	4	25.0	72	14	O89724	TAT PROTEIN (FRAGMENT)	1.12e+03
99	5	31.3	1120	10	O81509	T7M24.3 PROTEIN.	2.62e+01	172	4	25.0	72	14	O91891	TAT PROTEIN (FRAGMENT)	1.12e+03
100	5	31.3	1208	4	O75162	KIAA0675 PROTEIN.	2.62e+01	173	4	25.0	72	14	O91898	TAT PROTEIN (FRAGMENT)	1.12e+03
101	5	31.3	1253	5	O91601	ARYLPHORIN RECEPTOR.	2.62e+01	174	4	25.0	72	14	O75043	TAT.	1.12e+03
102	5	31.3	1253	5	O93809	ARYLPHORIN RECEPTOR.	2.62e+01	175	4	25.0	72	14	O75534	ENVELOPE GLYCOPROTEIN	1.12e+03
103	5	31.3	1280	5	O01779	COSMID W03F11.	2.62e+01	176	4	25.0	72	14	O90150	TAT PROTEIN (FRAGMENT)	1.12e+03
104	5	31.3	1417	10	O48697	F316.24 PROTEIN.	2.62e+01	177	4	25.0	72	14	O90150	TAT PROTEIN (FRAGMENT)	1.12e+03
105	5	31.3	1447	3	O09140	UDP-GLC:GLYCOPROTEIN G	2.62e+01	178	4	25.0	72	14	O86036	ISOLATE RL29 TAT (FRA	1.12e+03
106	5	31.3	1453	14	O91603	SPIKE PROTEIN.	2.62e+01	179	4	25.0	72	14	O86036	CLONE 800-106 TAT (FRA	1.12e+03
107	5	31.3	1513	5	O61567	RYANODINE RECEPTOR (FR	2.62e+01	180	4	25.0	72	14	O86001	CLONE 355-7 TAT (FRAGM	1.12e+03
108	5	31.3	1619	5	O77382	MAL3P6.13 PROTEIN.	2.62e+01	181	4	25.0	72	14	O75039	CLONE 355-6 TAT (FRAGM	1.12e+03
109	5	31.3	1690	3	O48635	PUTATIVE RNA BIOGENES	2.62e+01	182	4	25.0	72	14	O75038	TAT.	1.12e+03
110	5	31.3	1907	4	O45017	KIAA0299 (FRAGMENT).	2.62e+01	183	4	25.0	72	14	O75542	ENVELOPE GLYCOPROTEIN	1.12e+03
111	5	31.3	1927	5	O91133	SIMILAR TOS.	2.62e+01	184	4	25.0	72	14	O75525	ENVELOPE GLYCOPROTEIN	1.12e+03
112	5	31.3	2054	5	O17282	W09B6.1 PROTEIN.	2.62e+01	185	4	25.0	72	14	O74851	TAT (FRAGMENT).	1.12e+03
113	5	31.3	2388	11	O88197	BETA-SPECTRIN III.	2.62e+01	186	4	25.0	72	14	O86025	CLONE 760-167 TAT (FRA	1.12e+03
114	5	31.3	2390	4	O35020	BETA-SPECTRIN III.	2.62e+01	187	4	25.0	72	14	O86025	TAT PROTEIN (FRAGMENT)	1.12e+03
115	5	31.3	2743	4	O35162	POLYPOSIS LOCUS-ENCODE	2.62e+01	188	4	25.0	72	14	O90125	TAT PROTEIN (FRAGMENT)	1.12e+03
116	5	31.3	2829	13	P70039	ADENOMATOUS POLYPOSIS	2.62e+01	189	4	25.0	72	14	O90125	TAT PROTEIN (FRAGMENT)	1.12e+03
117	5	31.3	2842	11	P70478	APC PROTEIN.	2.62e+01	190	4	25.0	72	14	O86036	TAT PROTEIN (FRAGMENT)	1.12e+03
118	5	31.3	2844	4	O35163	POLYPOSIS LOCUS-ENCODE	2.62e+01	191	4	25.0	72	14	O86036	TAT PROTEIN (FRAGMENT)	1.12e+03
119	5	31.3	2845	11	O61315	ADENOMATOUS POLYPOSIS	2.62e+01	192	4	25.0	72	14	O90123	TAT PROTEIN (FRAGMENT)	1.12e+03
120	5	31.3	3085	4	O00183	BRC2 (FRAGMENT).	2.62e+01	193	4	25.0	72	14	O86038	TAT PROTEIN (FRAGMENT)	1.12e+03
121	5	31.3	3097	5	O82114	NAD(P)H-DEPENDENT GLUT	2.62e+01	194	4	25.0	72	14	O86038	CLONE 800-194 TAT (FRA	1.12e+03
122	5	31.3	3169	2	O82916	TOXIN B.	2.62e+01	195	4	25.0	72	14	O86039	CLONE 800-197 TAT (FRA	1.12e+03
123	5	31.3	3351	5	O34907	RETINOID- AND FATTY AC	2.62e+01	196	4	25.0	72	14	O75511	ENVELOPE GLYCOPROTEIN	1.12e+03
124	5	31.3	3488	5	O91257	SIMILAR TO C. ELEGANS	2.62e+01	197	4	25.0	72	14	O75588	ENVELOPE GLYCOPROTEIN	1.12e+03
125	5	31.3	4572	10	O40712	POLYPROTEIN.	2.62e+01	198	4	25.0	72	14	O75515	ENVELOPE GLYCOPROTEIN	1.12e+03
126	5	31.3	4981	5	O77372	MAL3P6.23 PROTEIN.	2.62e+01	199	4	25.0	72	14	O75516	ENVELOPE GLYCOPROTEIN	1.12e+03
127	5	31.3	5037	13	O91313	ALPHA-RYANODINE BINDIN	2.62e+01	200	4	25.0	72	14	O75524	ENVELOPE GLYCOPROTEIN	1.12e+03
128	5	31.3	6632	5	O17362	UNC-89.	2.62e+01	201	4	25.0	72	14	O75527	ENVELOPE GLYCOPROTEIN	1.12e+03
129	5	31.3	6642	5	O01761	C. ELEGANS UNC-89 (GB:	2.62e+01	202	4	25.0	72	14	O86031	CLONE 760-2 TAT (FRAGM	1.12e+03
130	4	25.0	15	2	O05991	SIMILAR TO BIN3 AND SI	1.12e+03	203	4	25.0	72	14	O86030	CLONE 760-193 TAT (FRA	1.12e+03
131	4	25.0	32	2	O05793	HYPOTHETICAL 3.8 KD PR	1.12e+03	204	4	25.0	72	14	O89484	CLONE 355-8 TAT (FRAGM	1.12e+03
132	4	25.0	34	2	P70846	ELECTRON TRANSFER FLAV	1.12e+03	205	4	25.0	72	14	O86020	CLONE 700-8 TAT (FRAGM	1.12e+03
133	4	25.0	43	2	OQ5847	HYPOTHETICAL PROTEIN I	1.12e+03	206	4	25.0	72	14	O69639	ISOLATE R249 TAT (FRA	1.12e+03
134	4	25.0	49	14	O55694	GP120 (FRAGMENT).	1.12e+03	207	4	25.0	72	14	O85979	CLONE 161-31 TAT (FRAG	1.12e+03
135	4	25.0	54	14	O90133	TAT PROTEIN.	1.12e+03	208	4	25.0	72	14	O85994	CLONE 355-29 TAT (FRAG	1.12e+03
136	4	25.0	60	14	O93032	HELICASE (FRAGMENT).	1.12e+03	209	4	25.0	72	14	O86057	CLONE 810-165 TAT (FRA	1.12e+03
137	4	25.0	60	14	O89633	TAT.	1.12e+03	210	4	25.0	72	14	O86058	CLONE 810-183 TAT (FRA	1.12e+03
138	4	25.0	61	1	O27420	HYPOTHETICAL 6.9 KD PR	1.12e+03	211	4	25.0	72	14	O86055	CLONE 810-134 TAT (FRA	1.12e+03
139	4	25.0	62	3	P79054	HYPOTHETICAL PROTEIN (	1.12e+03	212	4	25.0	72	14	O86017	CLONE 700-62 TAT (FRAG	1.12e+03
140	4	25.0	65	14	O75540	ENVELOPE GLYCOPROTEIN	1.12e+03	213	4	25.0	72	14	O86016	CLONE 700-61 TAT (FRAG	1.12e+03
141	4	25.0	66	9	O38293	E4.	1.12e+03	214	4	25.0	72	14	O85972	CLONE 161-15 TAT (FRAG	1.12e+03
142	4	25.0	68	14	O39997	CYTCHROME B (FRAGMENT	1.12e+03	215	4	25.0	72	14	O70504	ISOLATE DR105 TAT (FRA	1.12e+03
143	4	25.0	70	14	O89556	(STRAIN BANGLADESH-197	1.12e+03	216	4	25.0	72	14	O70536	ISOLATE DR7 TAT (FRAGM	1.12e+03
144	4	25.0	71	14	O91714	TAT PROTEIN (FRAGMENT)	1.12e+03	217	4	25.0	72	14	O90157	TAT PROTEIN (FRAGMENT)	1.12e+03
145	4	25.0	71	14	O91718	TAT PROTEIN (FRAGMENT)	1.12e+03	218	4	25.0	72	14	O90146	TAT PROTEIN (FRAGMENT)	1.12e+03
146	4	25.0	71	14	O91713	TAT PROTEIN (FRAGMENT)	1.12e+03	219	4	25.0	72	14	O90144	TAT PROTEIN (FRAGMENT)	1.12e+03
147	4	25.0	71	14	O91710	TAT PROTEIN (FRAGMENT)	1.12e+03	220	4	25.0	72	14	O90145	TAT PROTEIN (FRAGMENT)	1.12e+03
148	4	25.0	71	14	O91707	TAT PROTEIN (FRAGMENT)	1.12e+03	221	4	25.0	72	14	O89678	TAT.	1.12e+03
149	4	25.0	71	14	O91711	TAT PROTEIN (FRAGMENT)	1.12e+03	222	4	25.0	72	14	O74856	TAT (FRAGMENT).	1.12e+03
150	4	25.0	71	14	O91712	TAT PROTEIN (FRAGMENT)	1.12e+03	223	4	25.0	72	14	O74878	TAT (FRAGMENT).	1.12e+03
151	4	25.0	71	7	O31525	MHC CLASS IA (FRAGMENT	1.12e+03	224	4	25.0	72	14	O86037	CLONE 800-187 TAT (FRA	1.12e+03
152	4	25.0	71	14	O71898	CTRL 1, CLONE 1, ACCE	1.12e+03	225	4	25.0	72	14	O86040	CLONE 800-200 TAT (FRA	1.12e+03
153	4	25.0	71	14	O40225	TAT PROTEIN (FRAGMENT)	1.12e+03	226	4	25.0	72	14	O90136	TAT PROTEIN (FRAGMENT)	1.12e+03
154	4	25.0	71	14	O40230	TAT PROTEIN (FRAGMENT)	1.12e+03	227	4	25.0	72	14	O70498	ISOLATE D2-6 TAT (FRAG	1.12e+03
155	4	25.0	71	14	O40227	TAT PROTEIN (FRAGMENT)	1.12e+03	228	4	25.0	72	14	O70512	ISOLATE DR179 TAT (FRA	1.12e+03
156	4	25.0	71	14	O40226	TAT PROTEIN (FRAGMENT)	1.12e+03	229	4	25.0	72	14	O70500	ISOLATE D2-7 TAT (FRAG	1.12e+03
157	4	25.0	72	14	O89716	TAT PROTEIN (FRAGMENT)	1.12e+03	230	4	25.0	72	14	O70508	ISOLATE DR13 TAT (FRAG	1.12e+03
158	4	25.0	72	14	O89718	TAT PROTEIN (FRAGMENT)	1.12e+03	231	4	25.0	72	14	O90149	TAT PROTEIN (FRAGMENT)	1.12e+03
159	4	25.0	72	14	O89727	TAT PROTEIN (FRAGMENT)	1.12e+03	232	4	25.0	72	14	O90137	TAT PROTEIN (FRAGMENT)	1.12e+03
160	4	25.0	72	14	O89728	TAT PROTEIN (FRAGMENT)	1.12e+03	233	4	25.0	72	14	O74876	TAT (FRAGMENT).	1.12e+03
161	4	25.0	72	14	O89733	TAT PROTEIN (FRAGMENT)	1.12e+03	234	4	25.0	72	14	O74872	TAT (FRAGMENT).	1.12e+03
162	4	25.0	72	14	O89730	TAT PROTEIN (FRAGMENT)	1.12e+03	235	4	25.0	72	14	O85998	CLONE 355-40 TAT (FRAG	1.12e+03
163	4	25.0	72	14	O89729	TAT PROTEIN (FRAGMENT)	1.12e+03	236	4	25.0	72	14	O85999	CLONE 355-5 TAT (FRAGM	1.12e+03
164	4	25.0	72	14	O89913	TAT PROTEIN (FRAGMENT)	1.12e+03	237	4	25.0	72	14	O75585	ENVELOPE GLYCOPROTEIN	1.12e+03
165	4	25.0	72	14	O89731	TAT PROTEIN (FRAGMENT)	1.12e+03	238	4	25.0	72	14	O75575	ENVELOPE GLYCOPROTEIN	1.12e+03
166	4	25.0	72	14	O89725	TAT PROTEIN (FRAGMENT)	1.12e+03	239	4	25.0	72	14	O75574	ENVELOPE GLYCOPROTEIN	1.12e+03

240	4	25.0	72 14	Q75576	ENVELOPE GLYCOPROTEIN	1.12e+03	313	4	25.0	122 14	Q37881	C2 PROTEIN (FRAGMENT).	1.12e+03
241	4	25.0	72 14	Q75577	ENVELOPE GLYCOPROTEIN	1.12e+03	314	4	25.0	123 14	Q90857	GAG PROTEIN (FRAGMENT).	1.12e+03
242	4	25.0	72 14	Q75539	ENVELOPE GLYCOPROTEIN	1.12e+03	315	4	25.0	123 11	Q62196	SKELETAL MUSCLE RYANOD	1.12e+03
243	4	25.0	72 14	Q75560	ENVELOPE GLYCOPROTEIN	1.12e+03	316	4	25.0	124 14	Q90845	GAG PROTEIN (FRAGMENT)	1.12e+03
244	4	25.0	72 14	Q74862	TAT (FRAGMENT).	1.12e+03	317	4	25.0	124 7	Q31524	MHC CLASS IA (FRAGMENT	1.12e+03
245	4	25.0	72 14	Q74866	TAT (FRAGMENT).	1.12e+03	318	4	25.0	124 3	Q07438	ORF YDL071C.	1.12e+03
246	4	25.0	72 14	Q86024	CLONE 760-148 TAT (FRA	1.12e+03	319	4	25.0	124 5	Q21825	R07E5.5 PROTEIN.	1.12e+03
247	4	25.0	72 14	Q86028	CLONE 760-182 TAT (FRA	1.12e+03	320	4	25.0	125 14	Q90850	GAG PROTEIN (FRAGMENT)	1.12e+03
248	4	25.0	72 14	Q75517	ENVELOPE GLYCOPROTEIN	1.12e+03	321	4	25.0	125 14	Q90851	GAG PROTEIN (FRAGMENT)	1.12e+03
249	4	25.0	72 14	Q75514	ENVELOPE GLYCOPROTEIN	1.12e+03	322	4	25.0	125 14	Q91092	GAG PROTEIN (FRAGMENT)	1.12e+03
250	4	25.0	72 14	Q75046	TAT.	1.12e+03	323	4	25.0	125 14	Q73325	GAG PROTEIN (FRAGMENT)	1.12e+03
251	4	25.0	72 14	Q75047	TAT.	1.12e+03	324	4	25.0	125 14	Q73334	GAG PROTEIN (FRAGMENT)	1.12e+03
252	4	25.0	73 7	Q31519	MHC CLASS IA (FRAGMENT	1.12e+03	325	4	25.0	125 14	Q73310	GAG PROTEIN (FRAGMENT)	1.12e+03
253	4	25.0	73 1	Q21687	HYPOTHETICAL 8.3 KD PR	1.12e+03	326	4	25.0	125 14	Q73281	GAG PROTEIN (FRAGMENT)	1.12e+03
254	4	25.0	73 10	Q42894	1-AMINOCYCLOPROPANE-1-	1.12e+03	327	4	25.0	125 14	Q93219	GAG PROTEIN (FRAGMENT)	1.12e+03
255	4	25.0	73 7	Q31528	MHC CLASS IA (FRAGMENT	1.12e+03	328	4	25.0	125 14	Q73315	GAG PROTEIN (FRAGMENT)	1.12e+03
256	4	25.0	73 5	Q19279	SIMILARITY TO ZINC FIN	1.12e+03	329	4	25.0	125 14	Q73314	GAG PROTEIN (FRAGMENT)	1.12e+03
257	4	25.0	77 6	Q76322	MATRIX METALLOPROTEIN	1.12e+03	330	4	25.0	125 14	Q73273	GAG PROTEIN (FRAGMENT)	1.12e+03
258	4	25.0	78 5	Q25541	HYPOTHETICAL 9.3 KD PR	1.12e+03	331	4	25.0	125 14	Q73305	GAG PROTEIN (FRAGMENT)	1.12e+03
259	4	25.0	79 2	Q69880	HYPOTHETICAL 8.7 KD PR	1.12e+03	332	4	25.0	125 14	Q65171	PB125R.	1.12e+03
260	4	25.0	79 2	Q46229	RIBOSOMAL PROTEIN L22	1.12e+03	333	4	25.0	125 14	Q73294	GAG PROTEIN (FRAGMENT)	1.12e+03
261	4	25.0	79 14	Q69337	HYPOTHETICAL PROTEIN R	1.12e+03	334	4	25.0	125 14	Q73299	GAG PROTEIN (FRAGMENT)	1.12e+03
262	4	25.0	79 5	Q25520	SEGMENTATION PROTEIN (	1.12e+03	335	4	25.0	125 14	Q73330	GAG PROTEIN (FRAGMENT)	1.12e+03
263	4	25.0	80 8	Q33965	CYTOCHROME B (FRAGMENT	1.12e+03	336	4	25.0	125 14	Q73307	GAG PROTEIN (FRAGMENT)	1.12e+03
264	4	25.0	83 5	Q24662	SALSA (FRAGMENT).	1.12e+03	337	4	25.0	125 14	Q73331	GAG PROTEIN (FRAGMENT)	1.12e+03
265	4	25.0	86 14	Q74087	TAT.	1.12e+03	338	4	25.0	125 14	Q73260	GAG PROTEIN (FRAGMENT)	1.12e+03
266	4	25.0	86 14	Q90998	GP120 (FRAGMENT).	1.12e+03	339	4	25.0	127 14	Q73244	GAG PROTEIN (FRAGMENT)	1.12e+03
267	4	25.0	86 14	Q09778	TAT PROTEIN.	1.12e+03	340	4	25.0	128 14	Q90855	GAG PROTEIN (FRAGMENT)	1.12e+03
268	4	25.0	86 14	Q41769	HIV-1 ISOLATE 842R085	1.12e+03	341	4	25.0	129 14	Q90839	GAG PROTEIN (FRAGMENT)	1.12e+03
269	4	25.0	86 14	Q57293	TAT PROTEIN.	1.12e+03	342	4	25.0	129 2	Q03626	SKP PROTEIN (FRAGMENT)	1.12e+03
270	4	25.0	89 14	Q84146	B3L PROTEIN.	1.12e+03	343	4	25.0	130 14	Q90832	GAG PROTEIN (FRAGMENT)	1.12e+03
271	4	25.0	90 5	Q44091	VACUOLAR ATPASE 14KD S	1.12e+03	344	4	25.0	130 2	Q51201	CONSERVED HYPOTHETICAL	1.12e+03
272	4	25.0	91 9	Q80072	ORF 32.	1.12e+03	345	4	25.0	130 14	Q73277	GAG PROTEIN (FRAGMENT)	1.12e+03
273	4	25.0	93 1	Q74101	93AA LONG HYPOTHETICAL	1.12e+03	346	4	25.0	131 5	Q22641	T21C9.5 PROTEIN.	1.12e+03
274	4	25.0	98 14	Q41066	A584R PROTEIN.	1.12e+03	347	4	25.0	131 2	Q54897	M-LIKE PROTEIN PRECURS	1.12e+03
275	4	25.0	98 2	Q73171	HYPOTHETICAL 10.7 KD P	1.12e+03	348	4	25.0	132 14	Q91093	GAG PROTEIN (FRAGMENT)	1.12e+03
276	4	25.0	99 14	Q82995	TAT.	1.12e+03	349	4	25.0	132 14	Q73253	GAG PROTEIN (FRAGMENT)	1.12e+03
277	4	25.0	99 1	Q29734	HYPOTHETICAL 11.2 KD P	1.12e+03	350	4	25.0	132 2	Q07471	DNA GYRASE A (FRAGMENT	1.12e+03
278	4	25.0	101 14	Q92895	TAT PROTEIN.	1.12e+03	351	4	25.0	132 2	Q07477	DNA GYRASE A (FRAGMENT	1.12e+03
279	4	25.0	101 14	Q92899	TAT PROTEIN.	1.12e+03	352	4	25.0	133 14	Q73312	GAG PROTEIN (FRAGMENT)	1.12e+03
280	4	25.0	101 14	Q92885	TAT PROTEIN.	1.12e+03	353	4	25.0	133 14	Q73292	GAG PROTEIN (FRAGMENT)	1.12e+03
281	4	25.0	101 14	Q92893	TAT PROTEIN.	1.12e+03	354	4	25.0	133 14	Q85384	HOMOLOG OF VACCINIA VI	1.12e+03
282	4	25.0	101 14	Q93022	TAT PROTEIN.	1.12e+03	355	4	25.0	134 2	Q56573	SIMILAR TO VIBRIO PARA	1.12e+03
283	4	25.0	101 14	Q92902	TAT PROTEIN.	1.12e+03	356	4	25.0	135 4	Q14716	DNAP PROTEIN.	1.12e+03
284	4	25.0	101 14	Q89929	TAT PROTEIN.	1.12e+03	357	4	25.0	135 14	Q12279	TRAP PROTEIN.	1.12e+03
285	4	25.0	101 14	Q89943	TAT PROTEIN.	1.12e+03	358	4	25.0	135 14	Q88943	ORF C2.	1.12e+03
286	4	25.0	101 14	Q40219	TAT PROTEIN.	1.12e+03	359	4	25.0	137 14	Q55581	ORF26 PROTEIN.	1.12e+03
287	4	25.0	101 14	Q92880	TAT PROTEIN.	1.12e+03	360	4	25.0	138 10	Q48994	NBS-LRR TYPE RESISTANC	1.12e+03
288	4	25.0	101 14	Q93199	TAT PROTEIN.	1.12e+03	361	4	25.0	139 14	Q73255	GAG PROTEIN (FRAGMENT)	1.12e+03
289	4	25.0	101 14	Q90291	TAT PROTEIN.	1.12e+03	362	4	25.0	139 2	P71982	HYPOTHETICAL 16.2 KD P	1.12e+03
290	4	25.0	101 14	P89907	FROM AFRICAN GREEN MON	1.12e+03	363	4	25.0	141 14	Q73291	GAG PROTEIN (FRAGMENT)	1.12e+03
291	4	25.0	101 14	Q77688	TAT PROTEIN.	1.12e+03	364	4	25.0	141 2	Q57285	PESTICIN IMMUNITY PROT	1.12e+03
292	4	25.0	101 14	Q75758	TAT PROTEIN.	1.12e+03	365	4	25.0	141 5	Q00783	EBA-175 (FRAGMENT).	1.12e+03
293	4	25.0	101 14	Q74747	TAT PROTEIN.	1.12e+03	366	4	25.0	143 14	Q90835	GAG PROTEIN (FRAGMENT)	1.12e+03
294	4	25.0	101 14	Q56318	STRAIN SHIV-89.6, COMP	1.12e+03	367	4	25.0	143 14	Q11306	HOMOLOG OF VARIOLA B26	1.12e+03
295	4	25.0	101 14	Q93087	TAT PROTEIN.	1.12e+03	368	4	25.0	145 14	Q73284	GAG PROTEIN (FRAGMENT)	1.12e+03
296	4	25.0	101 8	Q79726	CYTOCHROME B (FRAGMENT	1.12e+03	369	4	25.0	146 14	Q73285	GAG PROTEIN (FRAGMENT)	1.12e+03
297	4	25.0	101 2	Q48022	LEX2A.	1.12e+03	370	4	25.0	146 11	Q88607	HYPOTHETICAL 15.2 KD P	1.12e+03
298	4	25.0	101 14	Q72990	IMMUNODEFICIENCY VIRUS	1.12e+03	371	4	25.0	148 3	P87290	HYPOTHETICAL 17.0 KD P	1.12e+03
299	4	25.0	101 14	Q74816	TAT PROTEIN.	1.12e+03	372	4	25.0	148 2	Q05738	HYPOTHETICAL 16.1 KD P	1.12e+03
300	4	25.0	101 14	Q80271	TAT.	1.12e+03	373	4	25.0	148 5	Q15650	PFEMP1 (FRAGMENT).	1.12e+03
301	4	25.0	101 14	Q55773	TAT PROTEIN.	1.12e+03	374	4	25.0	148 1	Q58732	HYPOTHETICAL PROTEIN M	1.12e+03
302	4	25.0	102 8	Q79725	CYTOCHROME B (FRAGMENT	1.12e+03	375	4	25.0	149 5	Q93846	K04G11.2 PROTEIN.	1.12e+03
303	4	25.0	103 10	P93086	PHYTOCHROME A-LIKE (FR	1.12e+03	376	4	25.0	152 8	Q21036	RIBOSOMAL PROTEIN S8.	1.12e+03
304	4	25.0	106 14	Q92297	TAT PROTEIN.	1.12e+03	377	4	25.0	152 2	Q45250	HYPOTHETICAL 17.7 KD P	1.12e+03
305	4	25.0	112 1	Q58272	112AA LONG HYPOTHETICA	1.12e+03	378	4	25.0	155 2	Q52410	ADENYLATE KINASE (FRAG	1.12e+03
306	4	25.0	115 14	Q90426	GP120 (FRAGMENT).	1.12e+03	379	4	25.0	155 2	Q52412	ADENYLATE KINASE (FRAG	1.12e+03
307	4	25.0	117 11	Q61879	NONMUSCLE MYOSIN HEAVY	1.12e+03	380	4	25.0	157 2	P74000	HYPOTHETICAL 18.9 KD P	1.12e+03
308	4	25.0	117 2	Q49314	RANDOM GENOMIC CLONE S	1.12e+03	381	4	25.0	157 2	P71457	PLASMID POG32 PLASMI	1.12e+03
309	4	25.0	119 8	Q34651	CYTOCHROME B (FRAGMENT	1.12e+03	382	4	25.0	158 14	Q89330	LECTIN HOMOLOG.	1.12e+03
310	4	25.0	119 8	Q34652	CYTOCHROME B (FRAGMENT	1.12e+03	383	4	25.0	158 2	P71105	HYPOTHETICAL PROTEIN (	1.12e+03
311	4	25.0	121 14	Q90844	GAG PROTEIN (FRAGMENT)	1.12e+03	384	4	25.0	158 10	Q40868	HEAT SHOCK-LIKE PROTEI	1.12e+03
312	4	25.0	121 5	Q24565	NUCLEOPLASMIN-LIKE PRO	1.12e+03	385	4	25.0	159 5	Q27666	SUPEROXIDE DISMUTASE (	1.12e+03

386	4	25.0	161	5	037294	R52.9 PROTEIN.	1.12e+03	459	25.0	258	2	P72768	HYPOTHETICAL 28.9 KD P	1.12e+03
387	4	25.0	162	5	015598	RIBOSOMAL PROTEIN S6 (	1.12e+03	460	25.0	258	2	056836	29KDA PROTEIN.	1.12e+03
388	4	25.0	163	14	089338	LECTIN HOMOLOG.	1.12e+03	461	25.0	261	2	054280	CTREF261 PROTEIN (FRAG	1.12e+03
389	4	25.0	166	5	093347	FLIC1.4 PROTEIN.	1.12e+03	462	25.0	261	2	087446	REPEAT MOTIF GENE A PR	1.12e+03
390	4	25.0	169	2	050570	GAM PROTEIN.	1.12e+03	463	25.0	262	14	092547	P31 PROTEIN.	1.12e+03
391	4	25.0	169	13	030801	ALPHA-1 COLLAGEN TYPE	1.12e+03	464	25.0	262	10	065118	ALCOHOL DEHYDROGENASE	1.12e+03
392	4	25.0	171	10	036374	PHYTOCHROME A (FRAGMEN	1.12e+03	465	25.0	262	10	065114	ALCOHOL DEHYDROGENASE	1.12e+03
393	4	25.0	174	5	015663	PFEMP1 (FRAGMENT).	1.12e+03	466	25.0	265	2	025017	HYPOTHETICAL 29.5 KD P	1.12e+03
394	4	25.0	176	11	060499	HOX-1.7 PROTEIN (GPK/5	1.12e+03	467	25.0	271	14	041036	A554L PROTEIN.	1.12e+03
395	4	25.0	177	10	048838	PUTATIVE UBIQUITIN CON	1.12e+03	468	25.0	271	2	050860	PFS PROTEIN (PFS).	1.12e+03
396	4	25.0	178	5	015667	PFEMP1 (FRAGMENT).	1.12e+03	469	25.0	272	14	089173	P3VC PROTEIN.	1.12e+03
397	4	25.0	184	5	035379	F19H8.3 PROTEIN.	1.12e+03	470	25.0	272	2	054375	DNAA (FRAGMENT).	1.12e+03
398	4	25.0	186	3	011118	P21 PROTEIN.	1.12e+03	471	25.0	272	10	065643	PUTATIVE MYC-TYPE TRAN	1.12e+03
399	4	25.0	187	5	062422	Y32F6B.3 PROTEIN.	1.12e+03	472	25.0	273	5	023134	SIMILARITY TO S. CEREV	1.12e+03
400	4	25.0	188	5	023412	COSMID ZK1193.	1.12e+03	473	25.0	277	9	038606	ORF2 (FRAGMENT).	1.12e+03
401	4	25.0	191	14	083020	ORF 3.	1.12e+03	474	25.0	279	14	088963	(WR) DNA-DEPENDENT RNA	1.12e+03
402	4	25.0	192	14	091081	VIF PROTEIN.	1.12e+03	475	25.0	280	2	085731	ROPB.	1.12e+03
403	4	25.0	193	2	073161	HYPOTHETICAL 22.1 KD P	1.12e+03	476	25.0	280	5	093239	C17G1.1 PROTEIN.	1.12e+03
404	4	25.0	192	2	094236	OUTER SURFACE PROTEIN	1.12e+03	477	25.0	281	2	034082	ORF51.	1.12e+03
405	4	25.0	194	10	048855	PROSYSTEMIN.	1.12e+03	478	25.0	283	14	071134	PUTATIVE MOVEMENT PROT	1.12e+03
406	4	25.0	194	5	045270	C27H2.4.	1.12e+03	479	25.0	283	11	035819	ZINC FINGER PROTEIN.	1.12e+03
407	4	25.0	195	10	050067	ALCOHOL DEHYDROGENASE	1.12e+03	480	25.0	288	2	008360	N-ACETYLNEURAMINATE LY	1.12e+03
408	4	25.0	196	3	040603	HYPOTHETICAL 22.0 KD P	1.12e+03	481	25.0	289	10	004271	CHITINASE.	1.12e+03
409	4	25.0	196	2	032812	GAMMA GLUTAMYL PHOSPHA	1.12e+03	482	25.0	290	4	099612	DNA-BINDING PROTEIN CP	1.12e+03
410	4	25.0	196	4	075672	HMGBCG PROTEIN (FRAGME	1.12e+03	483	25.0	290	2	074591	SHIKIMATE 5-DEHYDROGEN	1.12e+03
411	4	25.0	198	1	057724	198AA LONG HYPOTHETICA	1.12e+03	484	25.0	298	10	043808	ALCOHOL DEHYDROGENASE	1.12e+03
412	4	25.0	199	14	082449	RNA POLYMERASE (FRAGME	1.12e+03	485	25.0	300	8	020828	CYTOCHROME B (FRAGMENT	1.12e+03
413	4	25.0	199	4	035413	RYANODINE RECEPTOR TYP	1.12e+03	486	25.0	301	2	035756	SAM-P20 SERINE PROTEAS	1.12e+03
414	4	25.0	199	5	076321	SMALL GTPASE RACG.	1.12e+03	487	25.0	302	4	075840	UBIQUITOUS KRUPPEL LIK	1.12e+03
415	4	25.0	199	14	082448	RNA POLYMERASE (FRAGME	1.12e+03	488	25.0	304	14	090197	32 KDA ADSORPTION PROT	1.12e+03
416	4	25.0	201	3	032803	LPE20P.	1.12e+03	489	25.0	305	14	067859	DNA POLYMERASE (EC 2.7	1.12e+03
417	4	25.0	205	2	025493	CONSERVED HYPOTHETICAL	1.12e+03	490	25.0	306	3	006390	CHROMOSOME IV COSMID 9	1.12e+03
418	4	25.0	205	2	064666	HYPOTHETICAL 24.1 KD P	1.12e+03	491	25.0	308	14	098315	MC149R.	1.12e+03
419	4	25.0	205	5	045062	UBIQUITIN-CONJUGATING	1.12e+03	492	25.0	308	5	094277	SIMILAR TO SERINE/THRE	1.12e+03
420	4	25.0	208	8	021633	CYTOCHROME B (FRAGMENT	1.12e+03	493	25.0	308	5	025778	5' END (FRAGMENT).	1.12e+03
421	4	25.0	208	5	032930	COSMID C50E3.	1.12e+03	494	25.0	311	8	095798	CYTOCHROME C (FRAGMENT	1.12e+03
422	4	25.0	208	11	062572	(CLONE HORT13) (FRAGME	1.12e+03	495	25.0	311	8	095799	CYTOCHROME C (FRAGMENT	1.12e+03
423	4	25.0	211	10	049425	HYPOTHETICAL 24.1 KD P	1.12e+03	496	25.0	311	5	076712	F36H12.8 PROTEIN.	1.12e+03
424	4	25.0	214	2	086222	HYPOTHETICAL 25.1 KD P	1.12e+03	497	25.0	312	2	048836	PLASMD PGT2322, COMPLE	1.12e+03
425	4	25.0	215	5	045348	FLIP1.6 PROTEIN.	1.12e+03	498	25.0	313	10	080936	F13M22.18 PROTEIN.	1.12e+03
426	4	25.0	216	11	062568	(CLONE HORT4) (FRAGMEN	1.12e+03	499	25.0	319	1	029716	CONSERVED HYPOTHETICAL	1.12e+03
427	4	25.0	217	14	036648	REVERSE TRANSCRIPTASE	1.12e+03	500	25.0	319	2	054348	OUTER MEMBRANE PROTEIN	1.12e+03
428	4	25.0	217	2	067290	HYPOTHETICAL 25.6 KD P	1.12e+03	501	25.0	320	5	020225	COSMID F40F4.	1.12e+03
429	4	25.0	217	14	086801	CAPSID PROTEIN.	1.12e+03	502	25.0	321	5	016100	LOX20 HOMEODOMAIN PROT	1.12e+03
430	4	25.0	218	14	088619	COAT PROTEIN.	1.12e+03	503	25.0	323	9	064141	ATP/GTP BINDING PROTEI	1.12e+03
431	4	25.0	218	2	098813	HYPOTHETICAL 22.7 KD P	1.12e+03	504	25.0	323	2	054343	OUTER MEMBRANE PROTEIN	1.12e+03
432	4	25.0	218	14	066155	COAT PROTEIN.	1.12e+03	505	25.0	323	2	084700	HYPOTHETICAL 34.7 KD P	1.12e+03
433	4	25.0	220	14	036405	ORF55.	1.12e+03	506	25.0	324	5	062374	T09F5.10 PROTEIN.	1.12e+03
434	4	25.0	224	2	025733	REGULATORY COMPONENTS	1.12e+03	507	25.0	324	2	054346	OUTER MEMBRANE PROTEIN	1.12e+03
435	4	25.0	225	5	022314	COSMID T07E3.	1.12e+03	508	25.0	325	2	056880	V ANTIGEN (FRAGMENT).	1.12e+03
436	4	25.0	227	11	062457	(CLONE BALB8N).	1.12e+03	509	25.0	326	3	074553	CONSERVED HYPOTHETICAL	1.12e+03
437	4	25.0	227	11	062460	(CLONE BALB3N).	1.12e+03	510	25.0	326	10	093546	PEROXIDASE PRECURSOR (	1.12e+03
438	4	25.0	227	11	062459	(CLONE BALB11N).	1.12e+03	511	25.0	327	5	027276	MEC-5 PRECURSOR.	1.12e+03
439	4	25.0	234	5	021336	K08E4.6 PROTEIN.	1.12e+03	512	25.0	328	2	054347	OUTER MEMBRANE PROTEIN	1.12e+03
440	4	25.0	235	2	068188	INTEGRAL OUTER MEMBRAN	1.12e+03	513	25.0	328	2	087141	ORF139-26 PROTEIN.	1.12e+03
441	4	25.0	235	2	089330	ABC TRANSPORTER, ATP-B	1.12e+03	514	25.0	328	5	045313	C54E10.2 PROTEIN.	1.12e+03
442	4	25.0	236	2	071631	HYPOTHETICAL 25.8 KD P	1.12e+03	515	25.0	330	10	022649	ALCOHOL DEHYDROGENASE	1.12e+03
443	4	25.0	236	9	084059	HYPOTHETICAL 27.8 KD P	1.12e+03	516	25.0	331	13	079912	L-LACTATE DEHYDROGENAS	1.12e+03
444	4	25.0	236	2	073186	HYPOTHETICAL 26.4 KD P	1.12e+03	517	25.0	334	2	056896	V ANTIGEN (FRAGMENT).	1.12e+03
445	4	25.0	239	5	073334	ALCOHOL DEHYDROGENASE	1.12e+03	518	25.0	335	8	019920	YCF17 PROTEIN.	1.12e+03
446	4	25.0	245	10	023478	SIMILARITY TO NARINGEN	1.12e+03	519	25.0	335	5	076719	F36H12.3 PROTEIN.	1.12e+03
447	4	25.0	248	4	015812	SH3 DOMAIN-CONTAINING	1.12e+03	520	25.0	338	2	085692	ELECTRON-TRANSFERRING	1.12e+03
448	4	25.0	249	3	074385	HYPOTHETICAL 28.2 KD P	1.12e+03	521	25.0	339	4	060808	BR83387.1.	1.12e+03
449	4	25.0	249	2	035166	PUTATIVE NEURAMINILLAC	1.12e+03	522	25.0	339	5	025349	P4XP.	1.12e+03
450	4	25.0	251	2	033214	HYPOTHETICAL 26.8 KD P	1.12e+03	523	25.0	339	4	015080	P4OPHOX PROTEIN.	1.12e+03
451	4	25.0	251	3	060093	HYPOTHETICAL 29.2 KD P	1.12e+03	524	25.0	341	14	090760	ANK3.	1.12e+03
452	4	25.0	251	2	048792	TMS AND PRS GENES, PAR	1.12e+03	525	25.0	341	2	096310	ENDO-1,4-BETA-GLUCANAS	1.12e+03
453	4	25.0	252	10	043765	CHITINASE (EC 3.2.1.14	1.12e+03	526	25.0	342	5	077363	MAL3P4.5B PROTEIN.	1.12e+03
454	4	25.0	252	3	074510	VERY HYPOTHETICAL PROT	1.12e+03	527	25.0	342	2	068766	P2 PROTEIN HOMOLOG.	1.12e+03
455	4	25.0	256	5	095026	60 KD CHAPERONIN (PROT	1.12e+03	528	25.0	343	2	067374	ALCOHOL DEHYDROGENASE.	1.12e+03
456	4	25.0	257	14	089658	28KDA PROTEIN.	1.12e+03	529	25.0	343	2	054340	OUTER MEMBRANE PROTEIN	1.12e+03
457	4	25.0	257	5	076299	ACCESSORY GLAND-SPECIF	1.12e+03	530	25.0	344	5	091098	SIMILAR TO FAMILY 1 OF	1.12e+03
458	4	25.0	257	5	076298	ACCESSORY GLAND-SPECIF	1.12e+03	531	25.0	348	2	007616	HYPOTHETICAL 38.6 KD P	1.12e+03

532	4	25.0	349	1	026154	1.12e+03	605	419	5	021148	1.12e+03	SIMILAR TO DNA BINDING	1.12e+03
533	4	25.0	349	2	P72649	1.12e+03	606	421	14	041619	1.12e+03	ENVELOPE GLYCOPROTEIN	1.12e+03
534	4	25.0	351	2	086194	1.12e+03	607	423	3	Q04429	1.12e+03	D9719.31P	1.12e+03
535	4	25.0	355	5	077112	1.12e+03	608	425	3	Q13741	1.12e+03	CAMP-DEPENDENT PROTEIN	1.12e+03
536	4	25.0	355	3	014323	1.12e+03	609	426	10	004662	1.12e+03	DNAJ-LIKE PROTEIN.	1.12e+03
537	4	25.0	356	3	059694	1.12e+03	610	427	10	Q04587	1.12e+03	TAF-3.	1.12e+03
538	4	25.0	357	2	Q01452	1.12e+03	611	431	2	084367	1.12e+03	ASPARTOKINASE III.	1.12e+03
539	4	25.0	358	7	Q31012	1.12e+03	612	432	11	025479	1.12e+03	ANTI-CODON NUCLEASE MA	1.12e+03
540	4	25.0	360	14	041120	1.12e+03	613	432	11	035198	1.12e+03	PUTATIVE PHEROMONE REC	1.12e+03
541	4	25.0	360	11	Q61565	1.12e+03	614	436	10	022206	1.12e+03	T20B5.12 PROTEIN.	1.12e+03
542	4	25.0	362	14	092461	1.12e+03	615	437	13	013111	1.12e+03	AP-2 TRANSCRIPTION FAC	1.12e+03
543	4	25.0	362	14	067645	1.12e+03	616	437	5	Q20871	1.12e+03	HYPOPHETICAL 50.7 KD P	1.12e+03
544	4	25.0	362	5	017853	1.12e+03	617	439	10	Q04564	1.12e+03	T7N9.17.	1.12e+03
545	4	25.0	363	13	013130	1.12e+03	618	439	1	058431	1.12e+03	439AA LONG HYPOTHETICA	1.12e+03
546	4	25.0	363	2	P74321	1.12e+03	619	440	2	066886	1.12e+03	HYPOPHETICAL 52.6 KD P	1.12e+03
547	4	25.0	363	13	012971	1.12e+03	620	440	2	066886	1.12e+03	PUTATIVE HELICASE.	1.12e+03
548	4	25.0	366	7	046705	1.12e+03	621	442	14	055759	1.12e+03	SERINE PROTEASE (HTRA)	1.12e+03
549	4	25.0	367	7	046706	1.12e+03	622	443	2	025663	1.12e+03	444AA LONG HYPOTHETICA	1.12e+03
550	4	25.0	367	5	017230	1.12e+03	623	444	1	057949	1.12e+03	NEURAMINIDASE.	1.12e+03
551	4	25.0	369	1	P95879	1.12e+03	624	446	14	Q67027	1.12e+03	NEURAMINIDASE.	1.12e+03
552	4	25.0	373	3	013664	1.12e+03	625	446	14	Q67026	1.12e+03	CO4H5.5 PROTEIN.	1.12e+03
553	4	25.0	373	2	066701	1.12e+03	626	449	5	062027	1.12e+03	COUNTERPART OF HSV-1 G	1.12e+03
554	4	25.0	374	5	045770	1.12e+03	627	450	14	039292	1.12e+03	HYPOPHETICAL 51.4 KD P	1.12e+03
555	4	25.0	374	13	042866	1.12e+03	628	452	10	Q22232	1.12e+03	F26C24.3 PROTEIN.	1.12e+03
556	4	25.0	375	11	Q64673	1.12e+03	629	454	14	Q80022	1.12e+03	(HIVSBB, ISOLATE FROM	1.12e+03
557	4	25.0	375	11	Q64415	1.12e+03	630	454	2	050166	1.12e+03	MG2+ TRANSPORT PROTEIN	1.12e+03
558	4	25.0	375	11	Q64416	1.12e+03	631	455	2	Q53302	1.12e+03	CELLULASE.	1.12e+03
559	4	25.0	375	5	Q27904	1.12e+03	632	455	5	009945	1.12e+03	HYPOPHETICAL 52.4 KD P	1.12e+03
560	4	25.0	377	2	053452	1.12e+03	633	458	5	017607	1.12e+03	C27H6.2 PROTEIN.	1.12e+03
561	4	25.0	378	10	065459	1.12e+03	634	458	3	Q07684	1.12e+03	CHROMOSOME IV READING	1.12e+03
562	4	25.0	378	13	093363	1.12e+03	635	459	5	061268	1.12e+03	SHORT FORM OF NUCLEAR	1.12e+03
563	4	25.0	379	10	048868	1.12e+03	636	460	2	Q32738	1.12e+03	ADP-RIBOSYLTRANSFERASE	1.12e+03
564	4	25.0	379	10	023821	1.12e+03	637	461	5	044628	1.12e+03	F16B4.8 PROTEIN (FRAGM	1.12e+03
565	4	25.0	379	10	082430	1.12e+03	638	462	1	058999	1.12e+03	HYPOPHETICAL PROTEIN M	1.12e+03
566	4	25.0	380	8	048089	1.12e+03	639	466	5	015920	1.12e+03	PROTEIN PHOSPHATASE-BE	1.12e+03
567	4	25.0	380	10	Q43016	1.12e+03	640	467	14	056788	1.12e+03	N-ACETYLMURAMOYL-L-ALA	1.12e+03
568	4	25.0	380	10	Q43015	1.12e+03	641	469	14	Q82561	1.12e+03	NEURAMINIDASE.	1.12e+03
569	4	25.0	381	14	041941	1.12e+03	642	469	14	Q91744	1.12e+03	NEURAMINIDASE.	1.12e+03
570	4	25.0	382	5	Q94559	1.12e+03	643	469	14	092606	1.12e+03	NEURAMINIDASE.	1.12e+03
571	4	25.0	382	2	086319	1.12e+03	644	469	14	Q67055	1.12e+03	NEURAMINIDASE.	1.12e+03
572	4	25.0	383	5	Q26143	1.12e+03	645	469	14	Q91745	1.12e+03	NEURAMINIDASE.	1.12e+03
573	4	25.0	383	5	Q26144	1.12e+03	646	469	14	Q98815	1.12e+03	NEURAMINIDASE.	1.12e+03
574	4	25.0	383	5	Q26140	1.12e+03	647	469	14	Q98072	1.12e+03	NEURAMINIDASE.	1.12e+03
575	4	25.0	383	5	Q26150	1.12e+03	648	469	14	Q98073	1.12e+03	NEURAMINIDASE.	1.12e+03
576	4	25.0	385	1	058360	1.12e+03	649	469	14	Q67212	1.12e+03	NEURAMINIDASE. PRECURSO	1.12e+03
577	4	25.0	389	2	066180	1.12e+03	650	469	14	Q98064	1.12e+03	NEURAMINIDASE.	1.12e+03
578	4	25.0	389	5	Q94811	1.12e+03	651	469	14	Q98065	1.12e+03	NEURAMINIDASE.	1.12e+03
579	4	25.0	390	5	Q18401	1.12e+03	652	469	14	Q98066	1.12e+03	NEURAMINIDASE.	1.12e+03
580	4	25.0	390	2	083914	1.12e+03	653	469	14	Q67213	1.12e+03	NEURAMINIDASE. PRECURSO	1.12e+03
581	4	25.0	392	2	084399	1.12e+03	654	469	14	Q67214	1.12e+03	NEURAMINIDASE. PRECURSO	1.12e+03
582	4	25.0	393	5	001532	1.12e+03	655	470	11	062385	1.12e+03	INTERFERON-INDUCED PRO	1.12e+03
583	4	25.0	394	2	085335	1.12e+03	656	470	2	Q59556	1.12e+03	GLUTAMATE DECARBOXYLAS	1.12e+03
584	4	25.0	395	5	P90529	1.12e+03	657	476	14	Q83045	1.12e+03	RNA-DEPENDENT RNA POLY	1.12e+03
585	4	25.0	395	5	Q01531	1.12e+03	658	478	11	P97418	1.12e+03	LAN (LAN), SERINE THRE	1.12e+03
586	4	25.0	397	3	Q08144	1.12e+03	659	478	3	Q13862	1.12e+03	HYPOTHETICAL 51.8 KD 2	1.12e+03
587	4	25.0	401	5	061531	1.12e+03	660	478	10	081779	1.12e+03	HYPOPHETICAL 53.7 KD P	1.12e+03
588	4	25.0	401	10	Q48694	1.12e+03	661	480	4	075472	1.12e+03	TUMOROUS IMAGINAL DISC	1.12e+03
589	4	25.0	402	10	080595	1.12e+03	662	480	11	089110	1.12e+03	CASPASE-8.	1.12e+03
590	4	25.0	402	5	Q19217	1.12e+03	663	483	3	Q04408	1.12e+03	D9719.19P.	1.12e+03
591	4	25.0	404	14	Q36365	1.12e+03	664	484	5	Q44203	1.12e+03	TRACHEAE DEFECTIVE PRO	1.12e+03
592	4	25.0	405	2	Q51550	1.12e+03	665	484	6	Q28713	1.12e+03	SMOOTH MUSCLE MYOSIN H	1.12e+03
593	4	25.0	405	5	Q26268	1.12e+03	666	487	2	P76568	1.12e+03	FROM BASES 2613751 TO	1.12e+03
594	4	25.0	405	3	P87241	1.12e+03	667	487	13	Q73793	1.12e+03	NEUROFILAMENT MEDIUM S	1.12e+03
595	4	25.0	406	10	P93487	1.12e+03	668	488	2	Q51893	1.12e+03	HYPOPHETICAL 55.9 KD P	1.12e+03
596	4	25.0	407	4	Q60858	1.12e+03	669	491	10	Q22872	1.12e+03	T2P4.2 PROTEIN.	1.12e+03
597	4	25.0	408	2	085454	1.12e+03	670	494	14	Q84389	1.12e+03	CAPSID PROTEIN (FRAGME	1.12e+03
598	4	25.0	409	1	P96797	1.12e+03	671	494	14	Q84390	1.12e+03	CAPSID PROTEIN (FRAGME	1.12e+03
599	4	25.0	411	2	Q32743	1.12e+03	672	495	2	Q69080	1.12e+03	RNA POLYMERASE SUBUNIT	1.12e+03
600	4	25.0	412	2	P74401	1.12e+03	673	498	14	Q89489	1.12e+03	DNA COMPLETE GENOME.	1.12e+03
601	4	25.0	414	5	Q23119	1.12e+03	674	500	3	Q06682	1.12e+03	SIMILARITY NEAR C-TERM	1.12e+03
602	4	25.0	418	3	Q92236	1.12e+03	675	506	10	Q96560	1.12e+03	NADPH-FERRIHEMOPROTEIN	1.12e+03
603	4	25.0	418	3	Q92236	1.12e+03	676	507	8	Q47440	1.12e+03	F1 ATPASE ALPHA-SUBUNI	1.12e+03
604	4	25.0	419	2	Q30675	1.12e+03	677	508	5	045345	1.12e+03	F21H7.2 PROTEIN.	1.12e+03

678	4	25.0	510 11	054862	S2P.	1.12e+03	751	4	25.0	628 3	Q06344	CHROMOSOME IV COSMID 9	1.12e+03
679	4	25.0	515 5	044461	C09G12.7 PROTEIN.	1.12e+03	752	4	25.0	629 5	Q04567	H12119.4 PROTEIN.	1.12e+03
680	4	25.0	515 3	059923	SIR2.	1.12e+03	753	4	25.0	630 5	P91391	SIMILAR TO REVERSE TRA	1.12e+03
681	4	25.0	516 10	042701	CYTCHROME P450 (EC 1.	1.12e+03	754	4	25.0	631 5	Q44760	T12F5.5 PROTEIN.	1.12e+03
682	4	25.0	516 10	P93222	ADP-GLUCOSE SYNTHETASE	1.12e+03	755	4	25.0	631 5	Q48043	TRANSFERRIN BINDING PR	1.12e+03
683	4	25.0	519 14	024511	NONSTRUCTURAL PROTEIN	1.12e+03	756	4	25.0	642 5	Q12170	SIMILARITY OVER A SHOR	1.12e+03
684	4	25.0	520 10	022838	UNKNOWN PROTEIN.	1.12e+03	757	4	25.0	644 5	Q19636	F20C5.5 PROTEIN.	1.12e+03
685	4	25.0	522 4	015097	POLYADENYLATE BINDING	1.12e+03	758	4	25.0	649 13	P79984	HSC70 SHOCK PROTEIN.	1.12e+03
686	4	25.0	523 10	P93477	ADP-GLUCOSE SYNTHETASE	1.12e+03	759	4	25.0	650 13	Q73788	HEAT SHOCK PROTEIN 70.	1.12e+03
687	4	25.0	524 10	004924	ADP-GLUCOSE SYNTHETASE	1.12e+03	760	4	25.0	650 14	Q69489	GP105.	1.12e+03
688	4	25.0	524 10	042700	CYTCHROME P450 (EC 1.	1.12e+03	761	4	25.0	654 11	Q55152	CALPASTATIN.	1.12e+03
689	4	25.0	524 14	085411	GLYCOPROTEIN.	1.12e+03	762	4	25.0	658 14	Q41288	GAG PROTEIN.	1.12e+03
690	4	25.0	524 14	085410	GLYCOPROTEIN.	1.12e+03	763	4	25.0	658 14	Q90141	DNA POLYMERASE (FRAGME	1.12e+03
691	4	25.0	530 14	090713	VP2 PROTEIN (FRAGMENT)	1.12e+03	764	4	25.0	661 10	Q91360	ZEAXANTHIN EPOXIDASE.	1.12e+03
692	4	25.0	530 2	087219	TRSK PROTEIN (TRAK).	1.12e+03	765	4	25.0	663 3	Q43058	HNF-3/FORK HEAD TRANSC	1.12e+03
693	4	25.0	530 14	072726	VP2 (FRAGMENT).	1.12e+03	766	4	25.0	663 2	Q44860	P93 GENE (PARTIAL) (FR	1.12e+03
694	4	25.0	530 2	Q51530	PILIN TRANSCRIPTIONAL	1.12e+03	767	4	25.0	668 10	Q48532	COPPER AMINE OXIDASE P	1.12e+03
695	4	25.0	531 5	Q93651	F30F8.4 PROTEIN.	1.12e+03	768	4	25.0	668 5	Q50964	MCAC.	1.12e+03
696	4	25.0	531 5	062026	C04H5.3 PROTEIN.	1.12e+03	769	4	25.0	669 4	P81408	COTE1 PROTEIN.	1.12e+03
697	4	25.0	532 13	012989	CHORION PROTEIC COMPO	1.12e+03	770	4	25.0	670 11	Q01487	DNA-BINDING PROTEIN AT	1.12e+03
698	4	25.0	533 11	055221	PUTATIVE CD98 PROTEIN.	1.12e+03	771	4	25.0	671 10	Q39189	RNA HELICASE.	1.12e+03
699	4	25.0	534 10	080853	F7F1.9 PROTEIN.	1.12e+03	772	4	25.0	678 2	P71707	PROBABLE PENICILLIN-BI	1.12e+03
700	4	25.0	537 11	Q61098	INTERLEUKIN 1 RECEPT	1.12e+03	773	4	25.0	679 2	Q45012	P83/100 GENE.	1.12e+03
701	4	25.0	537 10	004319	POLY(A)-BINDING PROTEI	1.12e+03	774	4	25.0	682 3	Q13331	ACETOLACTATE SYNTHASE.	1.12e+03
702	4	25.0	539 11	070335	TYPE I ALPHA PHOSPHATI	1.12e+03	775	4	25.0	683 14	P89882	CAPSID PROTEIN VP1 (FR	1.12e+03
703	4	25.0	539 2	066192	60 KD CHAPERONIN (PROT	1.12e+03	776	4	25.0	683 5	Q09977	HYPOTHETICAL 76.4 KD P	1.12e+03
704	4	25.0	540 2	066212	60 KD CHAPERONIN (PROT	1.12e+03	777	4	25.0	683 14	Q84371	CAPSID PROTEIN VP1 (FR	1.12e+03
705	4	25.0	540 2	066214	60 KD CHAPERONIN (PROT	1.12e+03	778	4	25.0	684 1	Q27150	PHOSPHOENOLPYRUVATE SY	1.12e+03
706	4	25.0	540 2	Q92749	TYPE I PHOSPHATIDYLINO	1.12e+03	779	4	25.0	689 5	Q77377	MAL3P6.17 PROTEIN.	1.12e+03
707	4	25.0	541 10	048583	UNKNOWN PROTEIN.	1.12e+03	780	4	25.0	691 5	Q44741	F57B10.4 PROTEIN.	1.12e+03
708	4	25.0	541 2	066218	60 KD CHAPERONIN (PROT	1.12e+03	781	4	25.0	691 10	Q23295	HYPOTHETICAL 76.5 KD P	1.12e+03
709	4	25.0	541 2	066216	60 KD CHAPERONIN (PROT	1.12e+03	782	4	25.0	692 10	Q39035	NADPH-FERRIHEMOPROTEIN	1.12e+03
710	4	25.0	548 5	P91835	T2888.5 PROTEIN.	1.12e+03	783	4	25.0	695 9	Q00184	PUTATIVE ANTI-RECEPT	1.12e+03
711	4	25.0	549 13	Q90500	METALLOPROTEINASE (FRAG	1.12e+03	784	4	25.0	695 5	Q01856	SIMILARITY TO ACYLTRAN	1.12e+03
712	4	25.0	550 5	Q93764	F53C11.6 PROTEIN.	1.12e+03	785	4	25.0	696 5	Q17034	IMMUNE FACTOR.	1.12e+03
713	4	25.0	552 5	Q45812	T23G11.9 PROTEIN.	1.12e+03	786	4	25.0	699 2	Q31673	ATP-DEPENDENT CLP PROT	1.12e+03
714	4	25.0	556 10	082211	F6F22.31 PROTEIN (FRAG	1.12e+03	787	4	25.0	699 4	Q43263	RNA EDITING DEAMINASE	1.12e+03
715	4	25.0	556 10	Q38871	CALMODULIN-DOMAIN (PROT	1.12e+03	788	4	25.0	700 2	P72137	FLAGELLIN (FRAGMENT).	1.12e+03
716	4	25.0	558 10	Q38893	ZETA-CAROTENE DESATUR	1.12e+03	789	4	25.0	701 5	Q24928	IMMUNOGLOBULIN HEAVY C	1.12e+03
717	4	25.0	569 10	Q42982	4-COUMARATE:COA LIGASE	1.12e+03	790	4	25.0	702 2	Q22660	VGRE PROTEIN.	1.12e+03
718	4	25.0	574 14	Q89553	ORF18R.	1.12e+03	791	4	25.0	703 5	Q22559	T19B10.5 PROTEIN.	1.12e+03
719	4	25.0	575 2	Q01673	FLAGELLIN A.	1.12e+03	792	4	25.0	703 3	Q92227	HAEMAGGLUTININ.	1.12e+03
720	4	25.0	576 2	085179	MATRIX METALLOPROTEINA	1.12e+03	793	4	25.0	705 3	Q92227	PUTATIVE POLY(A)-BINDI	1.12e+03
721	4	25.0	579 5	Q61264	F26C24.6 PROTEIN.	1.12e+03	794	4	25.0	705 5	Q45884	W03B1.8 PROTEIN.	1.12e+03
722	4	25.0	580 10	080972	CAPSID PROTEIN 2.	1.12e+03	795	4	25.0	708 14	Q56960	PB1 PROTEIN.	1.12e+03
723	4	25.0	584 14	P89517	CAPSID PROTEIN VP2.	1.12e+03	796	4	25.0	712 5	Q20306	GCN20 PROTEIN HOMOLOG.	1.12e+03
724	4	25.0	584 14	Q96801	CAPSID PROTEIN (VP2).	1.12e+03	797	4	25.0	722 14	Q84394	DNA, COMPLETE GENOME S	1.12e+03
725	4	25.0	584 14	P87701	CAPSID PROTEIN (VP2).	1.12e+03	798	4	25.0	727 14	Q66206	ISOLATE CPV-128 COAT P	1.12e+03
726	4	25.0	584 14	P89947	CAPSID PROTEIN (VP2).	1.12e+03	799	4	25.0	727 14	Q83292	ISOLATE MEV-D COAT PRO	1.12e+03
727	4	25.0	584 14	P90438	CAPSID PROTEIN 2.	1.12e+03	800	4	25.0	727 5	Q19327	CODED FOR BY C. ELEGAN	1.12e+03
728	4	25.0	584 14	Q93187	CAPSID PROTEIN.	1.12e+03	801	4	25.0	727 14	Q67463	FELINE PANLEUKOPENIA I	1.12e+03
729	4	25.0	584 4	Q00178	PUTATIVE G-PROTEIN.	1.12e+03	802	4	25.0	731 14	Q87023	RNA POLYMERASE.	1.12e+03
730	4	25.0	584 14	Q84372	CAPSID PROTEIN VP2.	1.12e+03	803	4	25.0	738 6	P79390	PLATELET ENDOTHELIAL C	1.12e+03
731	4	25.0	584 14	Q70633	CAPSID PROTEIN.	1.12e+03	804	4	25.0	749 10	Q04387	MR 105,000 OUTER DYNEI	1.12e+03
732	4	25.0	584 14	P90357	CAPSID PROTEIN VP2.	1.12e+03	805	4	25.0	750 5	Q26191	MICRONEME PROTEIN-1 (F	1.12e+03
733	4	25.0	586 3	P79006	HNF-3/FORK HEAD TRANSC	1.12e+03	806	4	25.0	750 5	Q26187	MICRONEME PROTEIN-1 (F	1.12e+03
734	4	25.0	589 10	Q04599	A1C002N01.27.	1.12e+03	807	4	25.0	751 14	Q03725	COAT PROTEIN VP1 (CONT	1.12e+03
735	4	25.0	591 14	Q83391	GAG PRO POL POLYPROTEI	1.12e+03	808	4	25.0	751 14	Q03726	COAT PROTEIN VP1 (CONT	1.12e+03
736	4	25.0	591 4	Q15656	P68 TRK-T3 ONCOPROTEIN	1.12e+03	809	4	25.0	751 5	Q23395	2K1086.1 PROTEIN.	1.12e+03
737	4	25.0	591 10	Q39218	CHLOROPLAST RNA POLYME	1.12e+03	810	4	25.0	751 11	Q60816	ADAM 5 PROTEIN PRECURS	1.12e+03
738	4	25.0	598 5	Q16901	C3188.8 PROTEIN.	1.12e+03	811	4	25.0	769 4	Q00613	T3 RECEPTOR-ASSOCIATIN	1.12e+03
739	4	25.0	598 14	Q83353	ACETOHYDROXYACID SYNTH	1.12e+03	812	4	25.0	773 11	Q35601	FYN BINDING PROTEIN.	1.12e+03
740	4	25.0	605 2	Q83541	V-TYPE ATPASE, SUBUNIT	1.12e+03	813	4	25.0	774 5	Q17124	SPC3.	1.12e+03
741	4	25.0	609 3	Q89029	CHROMOSOME XVI READING	1.12e+03	814	4	25.0	778 4	Q12815	TASTIN.	1.12e+03
742	4	25.0	611 11	Q89020	ALPHA-ALBUMINE PROTEIN	1.12e+03	815	4	25.0	785 4	Q75843	GAMMA2-ADAPTIN.	1.12e+03
743	4	25.0	612 10	Q32318	HYPOTHETICAL 67.7 KD P	1.12e+03	816	4	25.0	789 14	Q88167	POLYMERASE.	1.12e+03
744	4	25.0	614 5	Q17499	B0395.3 PROTEIN.	1.12e+03	817	4	25.0	800 5	Q21145	KOZELL.1 PROTEIN.	1.12e+03
745	4	25.0	616 5	Q27361	ERYTHROCYTE BINDING AN	1.12e+03	818	4	25.0	808 5	Q94222	SIMILARITY TO SH3 DOMA	1.12e+03
746	4	25.0	616 5	Q27248	ERYTHROCYTE BINDING AN	1.12e+03	819	4	25.0	814 2	Q52984	HYPOTHETICAL 90.3 KD P	1.12e+03
747	4	25.0	618 14	Q86284	GAG POLYPROTEIN.	1.12e+03	820	4	25.0	815 2	Q47593	MANNOsylTRANSFERASE A.	1.12e+03
748	4	25.0	618 3	Q6817	SIMILARITY TO GLYCYL-TRNA	1.12e+03	821	4	25.0	817 4	Q14592	COMP_HUMAN.	1.12e+03
749	4	25.0	619 5	Q01759	SIMILARITY TO PHOSPHAT	1.12e+03	822	4	25.0	820 13	P79740	DEAD EYE PROTEIN.	1.12e+03
750	4	25.0	621 2	Q45169	P66 PROTEIN PRECURSOR.	1.12e+03	823	4	25.0	822 5	Q61674	FAINT SAUSAGE.	1.12e+03



824	4	25.0	825	3	074770	HYPOTHETICAL 92.5 KD P	1.12e+03	837	4	25.0	1156	14	093209	POL PROTEIN.	1.12e+03
825	4	25.0	830	14	083863	95.1KD PUTATIVE NONSTR	1.12e+03	838	4	25.0	1162	14	090022	RNA POLYMERASE SUBUNIT	1.12e+03
826	4	25.0	830	2	053129	PUTATIVE PENICILLIN-BI	1.12e+03	839	4	25.0	1175	5	024369	CALCIUM ACTIVATED POTA	1.12e+03
827	4	25.0	830	10	004074	STARCH BRANCHING ENZY	1.12e+03	900	4	25.0	1183	2	086064	CYTOTOXIN ASSOCIATED P	1.12e+03
828	4	25.0	832	14	066107	POLYMERASE.	1.12e+03	901	4	25.0	1183	3	008965	CHROMOSOME XVI READING	1.12e+03
829	4	25.0	833	14	087603	41KBP FRAGMENT FROM LE	1.12e+03	902	4	25.0	1199	14	083066	DNA-DEPENDENT RNA POLY	1.12e+03
830	4	25.0	834	14	086126	MEMBRANE GLYCOPROTEIN.	1.12e+03	903	4	25.0	1211	5	023260	ZC373.4 PROTEIN.	1.12e+03
831	4	25.0	835	14	057075	P95.	1.12e+03	904	4	25.0	1237	4	000114	HYPOTHETICAL HUMAN SER	1.12e+03
832	4	25.0	839	14	092446	P95+ACMPV ORF83.	1.12e+03	905	4	25.0	1238	11	088879	APOPTOTIC PROTEASE ACT	1.12e+03
833	4	25.0	842	2	053112	OLIGOPEPTIDE TRANSPO	1.12e+03	906	4	25.0	1246	14	036977	POL POLYPROTEIN (FRAG	1.12e+03
834	4	25.0	843	14	069605	DNA POLYMERASE (EC 2.7	1.12e+03	907	4	25.0	1263	5	009971	HYPOTHETICAL 145.0 KD	1.12e+03
835	4	25.0	843	2	034076	ORF45.	1.12e+03	908	4	25.0	1294	4	014529	KIA80293 (FRAGMENT).	1.12e+03
836	4	25.0	851	13	042507	THROMBOSPONDIN-3 (FR	1.12e+03	909	4	25.0	1323	2	087018	INACTIVE CYTOTOXIN.	1.12e+03
837	4	25.0	853	2	066463	HYPOTHETICAL 100.9 KD	1.12e+03	910	4	25.0	1325	5	077427	EG:115C2.3 PROTEIN.	1.12e+03
838	4	25.0	858	14	066144	CMV-SD 2A.	1.12e+03	911	4	25.0	1336	3	042803	DNA-C5-METHYLTRANSFER	1.12e+03
839	4	25.0	858	14	069535	STRUCTURAL PROTEIN.	1.12e+03	912	4	25.0	1353	3	074162	ICHI.	1.12e+03
840	4	25.0	859	14	092804	PR110.	1.12e+03	913	4	25.0	1390	5	077033	TRFA.	1.12e+03
841	4	25.0	860	5	027725	CYTIDINE TRIPHOSPHATE	1.12e+03	914	4	25.0	1398	11	005000	NUCLEAR RECEPTOR COACT	1.12e+03
842	4	25.0	860	3	036021	HYPOTHETICAL 98.0 KD P	1.12e+03	915	4	25.0	1403	11	088322	ENTACTIN-2.	1.12e+03
843	4	25.0	871	14	088155	ENVELOPE GLYCOPROTEIN.	1.12e+03	916	4	25.0	1406	10	004549	T7N9-2.	1.12e+03
844	4	25.0	872	2	047114	COOC PRECURSOR.	1.12e+03	917	4	25.0	1409	5	002626	REGULATOR OF PRESYNAPT	1.12e+03
845	4	25.0	876	10	081227	GAMMA-ADAPTIN 1.	1.12e+03	918	4	25.0	1412	14	088528	MAJOR CAPSID PROTEIN.	1.12e+03
846	4	25.0	893	14	069062	HERPESVIRUS TYPE 6 (HH	1.12e+03	919	4	25.0	1417	14	010688	RNA DEPENDENT RNA POLY	1.12e+03
847	4	25.0	907	2	069478	PYRUVATE DEHYDROGENASE	1.12e+03	920	4	25.0	1436	14	067726	NON-STRUCTURAL PROTEIN	1.12e+03
848	4	25.0	912	14	090278	ENVELOPE PROTEIN.	1.12e+03	921	4	25.0	1441	14	083001	POLYPROTEIN.	1.12e+03
849	4	25.0	913	13	013099	MIDDLE MOLECULAR WEIGH	1.12e+03	922	4	25.0	1441	14	040967	POLYPROTEIN.	1.12e+03
850	4	25.0	917	13	090883	LDL RECEPTOR HOMOLOGUE	1.12e+03	923	4	25.0	1441	5	026453	HLS-HOMELESS (FRAGMENT	1.12e+03
851	4	25.0	946	10	022015	HEP200 PROTEIN.	1.12e+03	924	4	25.0	1443	14	057107	POLYPROTEIN.	1.12e+03
852	4	25.0	946	10	042571	CHLOROPLAST ENVELOPE C	1.12e+03	925	4	25.0	1453	14	065984	SPIKE PROTEIN (FRAGMEN	1.12e+03
853	4	25.0	958	2	049931	PUTATIVE POLYKETIDE SY	1.12e+03	926	4	25.0	1458	5	008095	DYNEIN HEAVY CHAIN, CY	1.12e+03
854	4	25.0	971	2	051579	EXOEXOYRIBONUCLEASE V	1.12e+03	927	4	25.0	1471	5	046092	1-EVIDENCE-PREDICTED B	1.12e+03
855	4	25.0	977	2	025308	HYPOTHETICAL 110.9 KD	1.12e+03	928	4	25.0	1490	13	079922	MODIFICATION METHYLASE	1.12e+03
856	4	25.0	988	10	093661	DNA-DEPENDENT RNA POLY	1.12e+03	929	4	25.0	1490	3	074676	ABC TRANSPORTER.	1.12e+03
857	4	25.0	989	4	075448	THYROID HORMONE RECEPT	1.12e+03	930	4	25.0	1492	5	001319	FO2E3.4 PROTEIN.	1.12e+03
858	4	25.0	996	14	036638	DNA-DEPENDENT DNA POLY	1.12e+03	931	4	25.0	1492	2	068724	PHAGE LAMBDA HOST SPEC	1.12e+03
859	4	25.0	998	2	047262	HEMOLYSIN.	1.12e+03	932	4	25.0	1501	11	097393	P190-B.	1.12e+03
860	4	25.0	999	2	087296	HEMOGLOBIN BINDING PRO	1.12e+03	933	4	25.0	1532	5	077332	CHL3P3.16 PROTEIN.	1.12e+03
861	4	25.0	1001	2	025314	TYPE III RESTRICTION EN	1.12e+03	934	4	25.0	1547	3	099315	CHROMOSOME VII READING	1.12e+03
862	4	25.0	1004	14	098207	MC039L (EC 2.7.7.7).	1.12e+03	935	4	25.0	1548	5	091339	SIMILARITY TO HUMAN PR	1.12e+03
863	4	25.0	1004	1	028412	HYPOTHETICAL 115.0 KD	1.12e+03	936	4	25.0	1553	5	077385	MAL3P6.10 PROTEIN.	1.12e+03
864	4	25.0	1005	8	036613	CHLOROPLAST ORF313 AND	1.12e+03	937	4	25.0	1595	2	052373	FAMILY 10 XYLANASE (EC	1.12e+03
865	4	25.0	1005	1	058718	HYPOTHETICAL PROTEIN M	1.12e+03	938	4	25.0	1611	4	075089	AF-6.	1.12e+03
866	4	25.0	1012	4	051164	TRANSCRIPTION INTERMED	1.12e+03	939	4	25.0	1616	14	098745	180K PROTEIN.	1.12e+03
867	4	25.0	1016	5	061296	PERIOD PROTEIN.	1.12e+03	940	4	25.0	1616	14	093058	183K PEPTIDE.	1.12e+03
868	4	25.0	1021	2	046085	COLLAGENASE PRECURSOR.	1.12e+03	941	4	25.0	1616	14	098748	183K PEPTIDE.	1.12e+03
869	4	25.0	1026	14	036363	DNA POLYMERASE (EC 2.7	1.12e+03	942	4	25.0	1646	5	019338	F11A10.4 PROTEIN.	1.12e+03
870	4	25.0	1026	10	049529	PREDICTED PROTEIN.	1.12e+03	943	4	25.0	1653	2	076578	F1653.	1.12e+03
871	4	25.0	1028	5	076836	T19D12.4 PROTEIN.	1.12e+03	944	4	25.0	1744	5	022516	T14G8.1 (FRAGMENT).	1.12e+03
872	4	25.0	1037	2	033390	ALPHA-2,8-POLYSIALYLTR	1.12e+03	945	4	25.0	1761	10	048867	MODIFICATION METHYLASE	1.12e+03
873	4	25.0	1050	5	009655	HYPOTHETICAL 119.3 KD	1.12e+03	946	4	25.0	1777	14	089278	NONSTRUCTURAL PROTEIN	1.12e+03
874	4	25.0	1053	3	074425	3-HYDROXY-3-METHYLGLUT	1.12e+03	947	4	25.0	1779	14	098353	MURRAY VALLEY ENCEPHAL	1.12e+03
875	4	25.0	1055	2	025211	TYPE I RESTRICTION ENZ	1.12e+03	948	4	25.0	1787	5	018794	T14G8.1 PROTEIN.	1.12e+03
876	4	25.0	1061	5	020298	SIMILARITY TO PLASMODI	1.12e+03	949	4	25.0	1824	4	099976	MICROTUBULE-ASSOCIATED	1.12e+03
877	4	25.0	1067	14	089904	FROM AFRICAN GREEN MON	1.12e+03	950	4	25.0	1825	11	064715	MICROTUBULE-ASSOCIATED	1.12e+03
878	4	25.0	1069	14	069140	NUCLEAR ANTIGEN EBNA-3	1.12e+03	951	4	25.0	1828	11	063724	MICROTUBULE-ASSOCIATED	1.12e+03
879	4	25.0	1071	4	099433	CRM1 PROTEIN.	1.12e+03	952	4	25.0	1864	5	077368	CHL3P6.27 PROTEIN.	1.12e+03
880	4	25.0	1078	2	049529	PI20.	1.12e+03	953	4	25.0	1883	5	024376	CHROMODOMAIN-HELICASE-	1.12e+03
881	4	25.0	1078	5	027807	DYNEIN HEAVY CHAIN ISO	1.12e+03	954	4	25.0	1935	5	044934	MYOSIN HEAVY CHAIN ISO	1.12e+03
882	4	25.0	1083	5	024858	PYRIDINE NUCLEOTIDETRA	1.12e+03	955	4	25.0	1955	5	061308	MPRD.	1.12e+03
883	4	25.0	1084	5	024813	NICOTINAMIDE NUCLEOTID	1.12e+03	956	4	25.0	1979	11	088196	FATTY ACID SYNTHASE BE	1.12e+03
884	4	25.0	1086	2	069230	ENDO-1,4-BETA-XYLANASE	1.12e+03	957	4	25.0	1980	3	005747	SODIUM CHANNEL ALPHA-S	1.12e+03
885	4	25.0	1103	5	093804	EGL-15 PROTEIN.	1.12e+03	958	4	25.0	1993	5	090670	SODIUM CHANNEL ALPHA-S	1.12e+03
886	4	25.0	1104	10	049277	FIN21.17.	1.12e+03	959	4	25.0	2004	5	061222	ZK1005.1 PROTEIN (FRAG	1.12e+03
887	4	25.0	1106	5	018633	C44H9.4 PROTEIN.	1.12e+03	960	4	25.0	2017	14	088285	GAG-POL POLYPROTEIN.	1.12e+03
888	4	25.0	1116	5	018415	NUCLEAR PROTEIN SA.	1.12e+03	961	4	25.0	2098	5	025757	CTRP.	1.12e+03
889	4	25.0	1117	14	041340	130K PROTEIN.	1.12e+03	962	4	25.0	2109	14	065667	237K PROTEIN.	1.12e+03
890	4	25.0	1127	14	070852	POL POLYPROTEIN.	1.12e+03	963	4	25.0	2230	4	013439	TRANS-GOLGI P230.	1.12e+03
891	4	25.0	1130	8	035596	ORF-1130 (URFH PROTEIN	1.12e+03	964	4	25.0	2245	5	094378	ZC247.1 PROTEIN.	1.12e+03
892	4	25.0	1132	5	002432	CUP.	1.12e+03	965	4	25.0	2288	3	060033	ACETYL-COA CARBOXYLASE	1.12e+03
893	4	25.0	1134	1	030078	MOLYBDOTERIN OXIDORED	1.12e+03	966	4	25.0	2303	14	088595	THEILERS MURINE ENCEPH	1.12e+03
894	4	25.0	1140	5	023089	SIMILARITY TO CRM1 CHR	1.12e+03	967	4	25.0	2342	14	065980	RNA REPLICASE.	1.12e+03
895	4	25.0	1155	3	059742	HYPOTHETICAL 130.3 KD	1.12e+03	968	4	25.0	2344	3	074539	PUTATIVE SENSORY TRANSD	1.12e+03
896	4	25.0	1156	14	090458	REVERSE TRANSCRIPTASE.	1.12e+03	969	4	25.0	2457	5	077375	MAL3P6.20 PROTEIN.	1.12e+03

```

970 4 25.0 2498 14 Q86924 POLYPROTEIN 1. 1.12e+03
971 4 25.0 2513 14 Q90368 NONSTRUCTURAL POLYPROT 1.12e+03
972 4 25.0 2513 14 Q90370 NONSTRUCTURAL POLYPROT 1.12e+03
973 4 25.0 2541 5 Q13663 SIMILIARTO NON-MUSCLE M 1.12e+03
974 4 25.0 2657 5 Q77380 MAL3P6.15 PROTEIN. 1.12e+03
975 4 25.0 2713 4 Q14660 INOSITOL 1,4,5-TRISPHO 1.12e+03
976 4 25.0 2825 11 Q70465 DOC4. 1.12e+03
977 4 25.0 2864 14 Q69422 POLYPEPTIDE. 1.12e+03
978 4 25.0 2912 4 Q14054 COLLAGEN TYPE VII PREC 1.12e+03
979 4 25.0 3014 14 Q91936 POLYPROTEIN. 1.12e+03
980 4 25.0 3052 14 Q82933 JOHNSON GRASS MOSAIC V 1.12e+03
981 4 25.0 3056 14 Q65852 POLYPROTEIN. 1.12e+03
982 4 25.0 3086 14 Q89525 POLYPROTEIN. 1.12e+03
983 4 25.0 3225 4 Q14398 HUMAN GCP372. 1.12e+03
984 4 25.0 3343 11 Q35923 BREAST CANCER TYPE 2 S 1.12e+03
985 4 25.0 3432 14 Q90417 POLYPROTEIN. 1.12e+03
986 4 25.0 3432 14 Q82919 POLYPROTEIN. 1.12e+03
987 4 25.0 3432 14 Q82872 POLYPROTEIN. 1.12e+03
988 4 25.0 3432 14 Q90297 GP78 POLYPROTEIN. 1.12e+03
989 4 25.0 3432 14 Q92928 POLYPROTEIN PRECURSOR. 1.12e+03
990 4 25.0 3432 14 Q36765 POLYPROTEIN. 1.12e+03
991 4 25.0 3432 14 Q36764 POLYPROTEIN. 1.12e+03
992 4 25.0 3704 5 P91904 LAMININ ALPHA. 1.12e+03
993 4 25.0 3801 4 Q99698 LYOSOMAL TRAFFICKING 1.12e+03
994 4 25.0 4541 4 Q75592 PROTEIN ASSOCIATED WIT 1.12e+03
995 4 25.0 4823 13 Q93321 ALL-1 RELATED PROTEIN. 1.12e+03
996 4 25.0 4967 4 Q92736 RYANODINE RECEPTOR 2. 1.12e+03
997 4 25.0 5157 3 Q01135 PEPTIDE SYNTHETASE. 1.12e+03
998 4 25.0 5255 5 Q68006 BACTIRACIN SYNTHETASE 1.12e+03
999 4 25.0 7160 5 Q23551 ZK617.1B PROTEIN. 1.12e+03
1000 4 25.0 7962 4 Q10465 TITIN, SKELETAL MUSCLE 1.12e+03

```

## ALIGNMENTS

```

RESULT 1
ID O58162 PRELIMINARY; PRT; 440 AA.
AC O58162;
DT 01-AUG-1998 (TREMREL. 07, CREATED)
DT 01-AUG-1998 (TREMREL. 07, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)
DE 440AA LONG HYPOTHETICAL PROTEIN.
GN PH0425.
OS PYROCOCUS HORIKOSHII.
OC ARCHAEA; EURYARCHAEOTA; THERMOCOCCALES; THERMOCOCCACEAE; PYROCOCUS.
RN PRT SEQUENCE FROM N.A.
RP STRAIN=OT3;
RX MEDLINE: 98344137.
RA KAWABAYASHI Y., SAWADA M., HORIKAWA H., HAIKAWA Y., HINO Y.,
RA YAMAMOTO S., SEKINE M., BABA S., KOSUGI H., HOSIYAMA A., NAGAI Y.,
RA SAKAI M., OGURA K., OTUKA R., NAKAZAWA H., TAKAMIYA M., OHFUKU Y.,
RA FUNAHASHI T., TANAKA T., KUDOH Y., YAMAZAKI J., KUSHIDA N., OGUCHI A.,
RA AOKI K., NAKAMURA Y., ROBB T.F., HORIKOSHI K., MASUCHI Y., SHIZUYA H.,
RA KIKUCHI H.;
RT "Complete Sequence and Gene Organization of the Genome of a
RT Hyper-thermophilic Archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA RES. 5:53-76(1998).
DR EMBL: AP000002; D1030454; -.
SQ SEQUENCE 440 AA; 50609 MW; BBAEAB46 CRC32;

```

```

Query Match 37.5%; Score 6; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 3.59e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 366 INKDDV 371
|
|
|
|
|
Qy 2 INKDDV 7

```

```

RESULT 2
ID Q58327 PRELIMINARY; PRT; 574 AA.
AC Q58327;

```

```

DT 01-JUN-1998 (TREMREL. 06, CREATED)
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ0917.
GN MJ0917.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN PRT SEQUENCE FROM N.A.
RP MEDLINE: 96337999.
RX BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEORHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., SMITH H.O., WOESE C.R., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL SCIENCE 273:1058-1073(1996).
RN [2]
RP REVISIONS.
RX STRAIN=DSM 2661;
RX MEDLINE: 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEORHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., SMITH H.O., WOESE C.R., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: STRONG, IN THE C-TERMINAL, TO SYNECHOCYSTIS SP.
CC PCC 6803 SLR0400, E.COLI YFJB, H.INFLUENZAE HI0072 AND
CC M.TUBERCULOSIS H37Rv125.17.
DR EMBL: U67535; G2826350; -.
DR PFAM: PF00459; inositol_P; 2.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 574 AA; 64118 MW; 74047AAB CRC32;

```

```

Query Match 37.5%; Score 6; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 3.59e-01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 539 EINKDD 544
|
|
|
|
|
Qy 1 EINKDD 6

```

```

RESULT 3
ID Q89548 PRELIMINARY; PRT; 1968 AA.
AC Q89548;
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE ORF1 223K PROTEIN.
OS POTATO VIRUS M (PVM).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; CARLAVIRUS.
RN PRT SEQUENCE FROM N.A.
RP STRAIN=RUSSIAN WILD TYPE;
RA ZAVRIEV S.K.;
RL SUBMITTED (MAY-1990) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE OF 1-852 FROM N.A.
RC STRAIN=RUSSIAN WILD TYPE;
RX MEDLINE: 90301111.
RA RUPASOV V.V., LATHIGRAS R., KANYUKA K.V., RAMEZANI RAD M.,

```

RA LUKASHEVA L.I., ZAVRIEV S.K.:  
 RT "Nucleotide sequence and structural organization of the 3'-terminal  
 RL region of potato virus M.";  
 RM MOL. BIOL. (MOSK.) 24:448-459(1990).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RUSSIAN WILD TYPE;  
 RA ZAVRIEV S.K.;  
 RL SUBMITTED (OCT-1990) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RUSSIAN WILD TYPE;  
 RX MEDLINE: 92049299;  
 RA ZAVRIEV S.K., KANYUKA K.I., LEVAI K.E.:  
 RT "Complete nucleotide sequence of genomic RNA of the potato M-virus.";  
 RL MOL. BIOL. (MOSK.) 25:761-769(1991).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RUSSIAN WILD TYPE;  
 RA ZAVRIEV S.K.;  
 RL SUBMITTED (FEB-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
 [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 89293091.  
 RA RUPASOV V.V., MOROZOV S., YU KANYUKA K.V., ZAVRIEV S.K.:  
 RT "Partial nucleotide sequence of potato virus M RNA shows similarities  
 to protoviruses in gene arrangement and the encoded amino acid  
 sequences.";  
 RL J. GEN. VIROL. 70:1861-1869(1989).  
 DR EMBL: X53062; G61292; -;  
 DR EMBL: D14449; G222425; -;  
 SQ SEQUENCE 1968 AA; 223384 MW; 41BD5163 CRC32;

Query Match 37.5%; Score 6; DB 14; Length 1968;  
 Best Local Similarity 100.0%; Pred. No. 3.59e-01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1325 EINKDD 1330  
 |||||  
 QY 1 EINKDD 6

RESULT 4  
 ID O16004 PRELIMINARY; PRT; 2531 AA.  
 AC O16004;  
 DT 01-JAN-1998 (TREMREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)  
 DE NOTCH HOMOLOG.  
 OS LYTECHINUS VARIEGATUS (SEA URCHIN).  
 OC EUKARYOTA; METAZOA; ECHINODERMATA; ECHINOZOA; ECHINOIDEA; EUECHINOIDEA;  
 OC ECHINACEA; TENNOPLEUROIDA; TOXOPNEUSTIDAE; LYTECHINUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 97454256.  
 RA SHERWOOD D.R., MCCLAY D.R.:  
 RT "Identification and localization of a sea urchin Notch homologue:  
 insights into vegetal plate regionalization and Notch receptor  
 regulation.";  
 RL DEVELOPMENT 124:3363-3374(1997).  
 DR EMBL: AF000634; G2570351; -;  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 21.  
 DR PROSITE: PS01186; EGF\_2; 25.  
 DR PROSITE: PS01187; EGF\_CA; 20.  
 DR PFAM: PF00008; EGF; 34.  
 DR PFAM: PF00023; ank; 6.  
 DR PFAM: PF00066; notch; 3.  
 KW GLYCOPROTEIN; EGF-LIKE DOMAIN.  
 SQ SEQUENCE 2531 AA; 273982 MW; BB9C6F3D CRC32;

Query Match 37.5%; Score 6; DB 5; Length 2531;  
 Best Local Similarity 100.0%; Pred. No. 3.59e-01;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 EINKDD 360  
 |||||  
 QY 1 EINKDD 6

RESULT 5  
 ID Q38747 PRELIMINARY; PRT; 26 AA.  
 AC Q38747;  
 DT 01-NOV-1996 (TREMREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)  
 DE ASP13 MADS BOX (FRAGMENT)  
 OS ASPARAGUS OFFICINALIS (GARDEN ASPARAGUS).  
 OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; ASPARAGALES;  
 OC ASPARAGACEAE; ASPARAGUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ASP13;  
 RA MILLER H.G., KOCHER T.D., LOY B.:  
 RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: U07330; G497139; -;  
 DR PFAM: PF00319; transcript\_fact; 1.  
 DR MENDEL; 15132; ASPOf;MADS;mn15132.  
 FT NON\_TER 1 1  
 FT NON\_TER 26 26  
 SQ SEQUENCE 26 AA; 3095 MW; 6EF346D3 CRC32;

Query Match 31.3%; Score 5; DB 10; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 QVCFS 13  
 |||||  
 QY 11 QVCFS 15

RESULT 6  
 ID Q38753 PRELIMINARY; PRT; 26 AA.  
 AC Q38753;  
 DT 01-NOV-1996 (TREMREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)  
 DE ASP8 MADS BOX (FRAGMENT).  
 OS ASPARAGUS OFFICINALIS (GARDEN ASPARAGUS).  
 OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMERYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; ASPARAGALES;  
 OC ASPARAGACEAE; ASPARAGUS.  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ASP8;  
 RA MILLER H.G., KOCHER T.D., LOY B.:  
 RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: U07336; G497151; -;  
 DR PFAM: PF00319; transcript\_fact; 1.  
 DR MENDEL; 15136; ASPOf;MADS;mn15136.  
 FT NON\_TER 1 1  
 FT NON\_TER 26 26  
 SQ SEQUENCE 26 AA; 3066 MW; 557D2A15 CRC32;

Query Match 31.3%; Score 5; DB 10; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 9 QVCFS 13  
 |||||  
 QY 11 QVCFS 15

RESULT 7  
 ID Q67603 PRELIMINARY; PRT; 93 AA.  
 AC Q67603;  
 DT 01-NOV-1996 (TREMREL. 01, CREATED)

```

DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN ACL.
OS SQUASH LEAF CURL VIRUS.
OC VIRUSES; SS DNA VIRUSES; GEMINIVIRIDAE; BEGOMOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LOS MOCHIS 1.
RA LONIELO A.O., FORD R.E., SALINAS R.A., MORALES F.J., MAXWELL D.P.;
RL SUBMITTED (JAN-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR EMBL; L27272; G450637; -.
DR PFAM; PF00799; Geminivirus_AL1; 1.
FT NON_TER 1 93
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10550 MW; F93D645F CRC32;

Query Match 31.3%; Score 5; DB 14; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.62e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NKDDV 5
Qy 3 NKDDV 7

RESULT 8
ID P90283 PRELIMINARY; PRT; 102 AA.
AC P90283;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TAT PROTEIN.
GN TAT.
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1).
OC VIRUSES; RETROVIRUSES; RETROVIRIDAE; LENTIVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2321B;
RX MEDLINE; 9725063.
RA CHOI D.J., DUBE S., SPICER T.P., SLADE H.B., JENSEN F.C.,
RA POLESZ B.J.;
RT "HIV type 1 isolate Z321, the strain used to make a therapeutic HIV
RT type 1 immunogen, is intersubtype recombinant.";
RL AIDS RES. HUM. RETROVIRUSES 13:357-361(1997).
DR EMBL; U76035; G1857265; -.
DR PFAM; PF00539; Tat; 1.
SQ SEQUENCE 102 AA; 11645 MW; 28C8FC7F CRC32;

Query Match 31.3%; Score 5; DB 14; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.62e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 35 QVCFS 39
Qy 11 QVCFS 15

RESULT 9
ID Q53698 PRELIMINARY; PRT; 109 AA.
AC Q53698;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE BETA-LACTAMASE (EC 3.5.2.6) (FRAGMENT).
OS STAPHYLOCOCCUS AUREUS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC STAPHYLOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90095428.
RA EAST A.K., DYKE K.G.;
RT "Cloning and sequence determination of six Staphylococcus aureus

```

```

RT beta-lactamases and their expression in Escherichia coli and
RT Staphylococcus aureus.";
RL J. GEN. MICROBIOL. 135:1001-1015(1989).
CC -1- CATALYTIC ACTIVITY: A BETA-LACTAM + H(2)O - A SUBSTITUTED
CC BETA-AMINO ACID.
CC -1- SIMILARITY: THIS IS A CLASS-A BETA-LACTAMASE.
DR EMBL; M25256; G537340; -.
DR PROSITE; PS00146; BETA_LACTAMASE_A; 1.
DR PFAM; PF00144; beta-lactamase; 1.
KW HYDROLASE; ANTIBIOTIC RESISTANCE.
FT NON_TER 109 109
SQ SEQUENCE 109 AA; 12176 MW; 02D00570 CRC32;

Query Match 31.3%; Score 5; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.62e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 88 INKDD 92
Qy 2 INKDD 6

RESULT 10
ID Q25775 PRELIMINARY; PRT; 140 AA.
AC Q25775;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE DE CONSERVED HYPOTHETICAL PROTEIN.
GN HP1160.
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;
OC HELICOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695;
RX MEDLINE; 97394467.
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,
RA LEFLOS K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,
RA VENTER J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori [published erratum appears in Nature 1997 Sep
RT 25;389(6649):412].";
RL NATURE 388:539-547(1997).
DR EMBL; AE000822; G314318; -.
DR TIGR; HP1160; -.
DR PROSITE; PS01306; UPF0054; 1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 140 AA; 15646 MW; 779CB967 CRC32;

Query Match 31.3%; Score 5; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.62e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 EINKD 46
Qy 1 EINKD 5

RESULT 11
ID Q54273 PRELIMINARY; PRT; 141 AA.
AC Q54273;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF141 PROTEIN.
OS STAPHYLOCOCCUS SCIURI.

```

OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC STAPHYLOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K11 (792);  
RX MEDLINE: 98101461.  
RA WU S., DE LENCATRE H., TOMASZ A.;  
RT "Genetic organization of the mecA region in methicillin-susceptible  
RT and methicillin-resistant strains of Staphylococcus sciuri."  
RL J. BACTERIOL. 180:236-242(1998).  
DR EMBL: Y13094; E316582; -.  
SQ SEQUENCE 141 AA; 16365 MW; 4BAA6ACF CRC32;  
Query Match 31.3%; Score 5; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 6 QVCFS 10  
QY 11 QVCFS 15  
RESULT 12  
ID O05984 PRELIMINARY; PRT: 145 AA.  
AC O05984;  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 17.2 KD PROTEIN.  
OS STAPHYLOCOCCUS SCIURI.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC STAPHYLOCOCCUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC29062;  
RA WU S., PISCITELLI C., DE LENCATRE H., TOMASZ A.;  
RL MICROB. DRUG RES. 2:435-441(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC29062;  
RA WU S.;  
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC29062;  
RA WU S.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL: Y09223; E314910; -.  
DR PFAM: PF01047; Marf; 1.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 145 AA; 17166 MW; 27E3F17B CRC32;  
Query Match 31.3%; Score 5; DB 2; Length 145;  
Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 10 QVCFS 14  
QY 11 QVCFS 15  
RESULT 13  
ID Q59308 PRELIMINARY; PRT: 148 AA.  
AC Q59308;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
DE FOLYLPOLYGLUTAMATE SYNTHASE (EC 6.3.2.17)  
DE (FOLYLPOLY-GAMMA-GLUTAMATE SYNTHETASE) (FPGS) (FRAGMENT).  
GN FOLC.  
OS CLOSTRIDIUM PERFRINGENS.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;  
OC CLOSTRIDIUM.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CPN50;  
RA KATAYAMA S., DUPUY B., GARNIER T., COLE S.T.;  
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: CONVERSION OF FOLATES TO POLYGLUTAMATE DERIVATIVES.  
CC -!- CATALYTIC ACTIVITY: ATP + TETRAHYDROFOLYL-[GLU][N] + L-GLUTAMATE  
CC -!- - ADP + ORTHOPHOSPHATE + TETRAHYDROFOLYL-[GLU][N+1].  
CC -!- SIMILARITY: TO OTHER FOLYLPOLYGLUTAMATE SYNTHASE.  
DR EMBL: X86524; E149544; -.  
DR PROSITE: PS01011; FOLYLPOLYGLU SYNT.1; 1.  
DR PROSITE: PS01012; FOLYLPOLYGLU SYNT.2; 1.  
KW LIGASE; ONE-CARBON METABOLISM; ATP-BINDING.  
FT NON\_TER 1  
SQ SEQUENCE 148 AA; 16421 MW; C5B83C1F CRC32;  
Query Match 31.3%; Score 5; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 69 INKDD 73  
QY 2 INKDD 6  
RESULT 14  
ID Q26067 PRELIMINARY; PRT: 157 AA.  
AC Q26067;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE MYOSIN REGULATORY LIGHT CHAIN.  
OS PLACOPECTEN MAGELLANICUS (SEA SCALLOP).  
OC EUKARYOTA; METAZOA; MOLLUSCA; BIVALVIA; PTERIOMORPHIA; OSTREOIDA;  
OC PECTINIDAE; PLACOPECTEN.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PERREAU-MICALE C., JANCOS A., SZENT-GYORGYI A.G.;  
RL J. MUSCLE RES. CELL. MOTIL. 0:0-0(0).  
DR EMBL: U59291; G1389845; -.  
DR PFAM: PF00036; efhand; 1.  
SQ SEQUENCE 157 AA; 17691 MW; 09657840 CRC32;  
Query Match 31.3%; Score 5; DB 5; Length 157;  
Best Local Similarity 100.0%; Pred. No. 2.62e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 36 INKDD 40  
QY 2 INKDD 6  
RESULT 15  
ID O30452 PRELIMINARY; PRT: 173 AA.  
AC O30452;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 19.4 KD PROTEIN.  
OS CLOSTRIDIUM PASTEURIANUM.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;  
OC CLOSTRIDIUM.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DSM 525;  
RX MEDLINE: 97457194.  
RA LUPERS F., SEYFRIED M., DANIEL R., GOTTSCHALK G.;  
RT "Glycerol conversion to 1,3-propanediol by Clostridium pasteurianum:  
RT cloning and expression of the gene encoding 1,3-propanediol  
RT dehydrogenase."  
RL FEMS MICROBIOL. LETT. 154:337-345(1997).  
DR EMBL: AF006034; G2393885; -.  
KW HYPOTHETICAL PROTEIN.

SQ SEQUENCE 173 AA: 19439 MW: 04FDE4F3 CRC32:  
 Query Match 31.3%; Score 5; DB 2; Length 173;  
 Best Local Similarity 100.0%; Pred.No. 2.62e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 21 INKDD 25  
 |||||  
 Qy 2 INKDD 6

Search completed: Sat Aug 28 15:12:07 1999  
 Job time : 48 secs.

\*\*\*\*\*

W P S R L (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:04:04 1999; MasPar time 6.67 Seconds  
Tabular output not generated. 60.601 Million cell updates/sec

Title: >US-09-049-696-42  
Description: (1-19) from US09049696.pap  
Perfect Score: 144  
Sequence: 1 YNNDKFYLSNGRIQAVRC 19

Scoring table: PAM 150  
Gap 15

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35

1:part1 2:part3 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 20.741; Variance 64.639; scale 0.321

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	63	43.8	947	15	R81953	Rice pyruvate orthophosphate dikinase (PPDK)
2	63	43.8	947	18	R90924	Corn derived pyruvate orthophosphate dikinase (PPDK)
3	63	43.8	947	14	R76708	Recombinant cold-resistance pyruvate orthophosphate dikinase (PPDK)
4	63	43.8	953	18	R90923	F. biguttatus pyruvate orthophosphate dikinase (PPDK)
5	63	43.8	953	14	R76707	Recombinant cold-resistance pyruvate orthophosphate dikinase (PPDK)
6	63	43.8	955	18	R90927	F. browningii cold-resistance pyruvate orthophosphate dikinase (PPDK)
7	63	43.8	955	14	R76711	Recombinant cold-resistance pyruvate orthophosphate dikinase (PPDK)
8	61	42.4	350	31	W56265	Mouse recombinase mur
9	59	41.0	25	35	W79304	Staphylococcus aureus
10	59	41.0	446	12	R66448	C. jejuni hippuricase
11	59	41.0	621	27	W36451	Human tyrosine phosphatase
12	58	40.3	90	24	W21828	Rat RII.Balpa-sd c1a
13	58	40.3	258	3	R13117	Cholera toxin A1 frag
14	58	40.3	258	26	W25785	Vibrio cholerae toxin
15	58	40.3	1169	18	R96126	Bacillus thuringiensis
16	57	39.6	46	15	R81967	ORF encoded by exon 1

17	57	39.6	413	14	R85327	IS1096 transposase.	9.87e+01
18	56	38.9	200	34	W38593	S. pneumoniae ATP-dep	1.24e+02
19	56	38.9	234	36	W74896	Human secreted protei	1.24e+02
20	56	38.9	270	22	W06484	Carbonic anhydrase do	1.24e+02
21	56	38.9	496	39	W86161	Insect GABA receptor	1.24e+02
22	56	38.9	496	15	R89336	GABA receptor subunit	1.24e+02
23	56	38.9	668	1	P81186	Sequence encoded by n	1.24e+02
24	56	38.9	668	1	P82924	Cowpox virus protein	1.24e+02
25	56	38.9	2308	10	R52580	RTPP-beta amino acid	1.24e+02
26	56	38.9	2308	11	R57902	Human RPP-beta.	1.24e+02
27	55	38.2	350	31	W56284	Human recombinase hsr	1.55e+02
28	55	38.2	479	25	W16794	Aspartase-470.	1.55e+02
29	55	38.2	481	22	W20344	H. pylori putative me	1.55e+02
30	55	38.2	595	38	W87487	S. cerevisiae TIH1 po	1.55e+02
31	55	38.2	595	14	R81314	Yeast TIH1.	1.55e+02
32	55	38.2	595	31	W54167	S. cerevisiae TIH1 pro	1.55e+02
33	55	38.2	702	3	P60230	Dihydroxyacetone-synt	1.55e+02
34	55	38.2	702	2	R11607	Recombinant dihydroxy	1.55e+02
35	55	38.2	957	22	W20841	H. pylori transembra	1.55e+02
36	54	37.5	548	27	W27377	L. cruciata luciferas	1.55e+02
37	54	37.5	548	3	R14121	Luciferase mutant Tyr	1.55e+02
38	54	37.5	595	39	W89184	Os2/cbfal native pol	1.94e+02
39	54	37.5	857	17	R95287	Nasturtium xyloglucan	1.94e+02
40	54	37.5	1206	3	P60679	Sequence of Riff Vail	1.94e+02
41	53	36.8	119	24	W21827	Rat RII.Balpal class	2.43e+02
42	53	36.8	165	17	R91063	Turkey prepro-VIP.	2.43e+02
43	53	36.8	173	8	R51078	Polyptide encoded b	2.43e+02
44	53	36.8	218	23	W24387	Lymantria dispar nucl	2.43e+02
45	53	36.8	1454	9	R42470	Feline infectious per	2.43e+02

## ALIGNMENTS

RESULT 1  
ID R81953 standard; Protein; 947 AA.  
AC R81953;

DE Rice pyruvate orthophosphate dikinase (exons 2-21).  
KW Rice; promoter; exon; tissue-specific; season-specific; PPDK;  
KW pyruvate orthophosphate dikinase; C4 photosynthesis; C3 plant.

OS Oryza sativa L.

FH Key Location/Qualifiers

FT peptide 1..71 /label=transit\_peptide

PN J07184657-A.

PD 25-JUL-1995.

PF 28-DEC-1993; 335671.

PR 28-DEC-1993; JP-335671.

PA (NITA-) JAPAN TAFU GURASU KK.

DR WPI: 95-287968/38.

PT Pyruvate orthophosphate dikinase (C4 photosynthesis-related) gene  
and promoter in rice - shows tissue-specific and seasonal specific  
expression.

PS Claim 1; Page 12-14; 26pp; Japanese.

CC The sequence is the rice pyruvate orthophosphate dikinase (PPDK)  
CC encoded by exons 2 to 21 of the gene shown in 098241-61. The gene  
CC shows tissue-specific (in rice leaf and rice arista) and seasonal  
CC specific expression. The PPDK gene may give a C4 photosynthesis  
CC function to a C3 plant.

SQ Sequence 947 AA;

Query Match 43.8%; Score 63; DB 15; Length 947;  
Best Local Similarity 46.7%; Pred. No. 2.45e+01;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 637 ehmfassderikavr 651

QY 4 DEKFYLSNGRIQAVR 18

RESULT 2

ID R90924 standard; Protein; 947 AA.  
AC R90924;

```

RESULT      4
ID   R30923 standard; Protein; 953 AA.
AC   R0923;
DT   26-SEP-1996 (first entry)
DE   F. bidentis pyruvate phosphate dikinase.
KW   Wild type; pyruvate phosphate dikinase; PPK; cold-resistant;
KW   F. brownii; F. bidentis; maize; cold resistance.
OS   Flaveria bidentis.
PN   WO9604369-A1.
PD   15-FEB-1996.
PF   30-MAY-1995; J01040.
PR   23-JUL-1994; JP-197780.
PR   01-DEC-1994; WO-J02022.
PA   (NISR ) JAPAN TOBACCO INC.
PI   Burnell NJ, Ohta S, Usami S;
DR   WPI: 96-129386/13.
DR   N-PSDB; T12672.
PT   Recombinant cold-resistant pyruvate phosphate di:kinase - imparts
PT   improved cold tolerance to transgenic plants which express the
PT   enzyme
PS   Claim 1; Page 23-29; 70pp; Japanese.
CC   The sequences given in R30923-26 represent wild type pyruvate phosphate
CC   dikinase (PPDK) enzymes from various organisms. These sequences
CC   were used in the design of a novel protein having cold-resistant
CC   PPK activity (see also R90927). The cold-resistant PPK is based
CC   primarily on the PPK sequence of F. brownii and comprises an
CC   addition, deletion or substitution of one or more amino acids in the
CC   C-terminal portion, esp. within the region residues 832-955. The
CC   cold-resistant PPK may also be based on the F. bidentis PPK
CC   sequence having Pro at position 869 and/or Leu at 885 and Val at 952.
CC   Plants, e.g. maize, transformed with DNA encoding the cold-resistant
CC   PPK, have improved cold resistance.
SQ   Sequence 953 AA;

Query Match      43.8%; Score 63; DB 18; Length 953;
Best Local Similarity 46.7%; Pred. No. 2.45e+01;
Matches      7; Conservative      4; Mismatches      0; Gaps      0;

Db      643 ehmfassderikavr 657
QY      4 DEXFYLNGRIQAVR 18
      : : l: l: l: l: l:

RESULT      5
ID   R76707 standard; Protein; 953 AA.
AC   R76707;
DT   25-JAN-1996 (first entry)
DE   Recombinant cold-resistant pyruvate phosphate dikinase.
KW   Recombinant pyruvate phosphate dikinase; cold-resistant;
KW   transformed plants; maize; deccan grass; millet; sorghum.
OS   Flaveria bidentis.
PN   WO9515385-A1.
PD   08-JUN-1995.
PF   01-DEC-1994; J02022.
PR   03-DEC-1993; AU-052275.
PR   29-JUL-1994; JP-197780.
PA   (NISR ) JAPAN TOBACCO INC.
PI   Burnell NJ, Ohta S, Usami S;
DR   WPI: 95-215265/28.
DR   N-PSDB; Q94641.
PT   Recombinant cold-resistant pyruvate phosphate di:kinase - and DNA
PT   coding for it which can be used to impart cold resistance to plants.
PS   Claim 1; Pages 23-29; 70pp; Japanese.
CC   Q94641 encodes R76707 a recombinant cold-resistant pyruvate phosphate
CC   dikinase (PPDK). The recombinant PPK can be used in the prodn. of
CC   transformed cold-resistant plants, e.g. maize, deccan grass, millet
CC   or sorghum.
SQ   Sequence 953 AA;

Query Match      43.8%; Score 63; DB 14; Length 953;
Best Local Similarity 46.7%; Pred. No. 2.45e+01;

```



Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 643 ehmfassderikavr 657

QY 4 DEKFYLSNGRIQAVR 18

# RESULT 6

ID R0927 standard; Protein; 955 AA.

AC R0927; 955 AA (first entry)

DE F. brownii cold resistant pyruvate phosphate dikinase.

KW Wild type; pyruvate phosphate dikinase; PPK; cold-resistant;

OS Flaveria brownii.

PN WO9604369-A1.

PD 15-FEB-1996.

PF 30-MAY-1995; J01040.

PR 29-JUL-1994; JP-197780.

PR 01-DEC-1994; WO-J02022.

PA (NLSB ) JAPAN TOBACCO INC.

PI Burnell NJ, Ohta S, Usami S;

DR WPI; 96-129386/13.

DR N-PSDB: T12676.

PT Recombinant cold-resistant pyruvate phosphate dikinase - imparts

PT improved cold tolerance to transgenic plants which express the

PT enzyme

PS Claim 2; Page 47-54; 70pp; Japanese.

CC This sequence represents a mutant pyruvate phosphate dikinase (PPDK)

CC enzyme. This sequence is based on the sequences given in R0923-26

CC which represent wild type PPK enzymes from various organisms. This

CC mutated sequence represents a novel protein having cold-resistant PPK

CC activity. The cold-resistant PPK is based primarily on the PPK

CC sequence of F. brownii and comprises an addition, deletion or

CC substitution of one or more amino acids in the C-terminal portion,

CC esp. within the region residues 832-955. The cold-resistant PPK

CC may also be based on the F. bidentis PPK sequence having Pro at

CC position 869 and/or Leu at 885 and Val at 952. Plants, e.g. maize,

CC transformed with DNA encoding the cold-resistant PPK, have improved

CC cold resistance.

SQ Sequence 955 AA;

Query Match 43.8%; Score 63; DB 18; Length 955;

Best Local Similarity 46.7%; Pred. No. 2.45e+01;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 645 ehmfassderikavr 659

QY 4 DEKFYLSNGRIQAVR 18

# RESULT 7

ID R76711 standard; Protein; 955 AA.

AC R76711; 955 AA (first entry)

DE Recombinant cold-resistant pyruvate phosphate dikinase.

KW Recombinant pyruvate phosphate dikinase; cold-resistant;

OS Flaveria brownii.

PN WO9515385-A1.

PD 08-JUN-1995.

PF 01-DEC-1994; J02022.

PR 03-DEC-1993; AU-052275.

PR 29-JUL-1994; JP-197780.

PA (NLSB ) JAPAN TOBACCO INC.

PI Burnell NJ, Ohta S, Usami S;

DR WPI; 95-215265/28.

DR N-PSDB: 094645.

PT Recombinant cold-resistant pyruvate phosphate dikinase - and DNA

PT coding for it which can be used to impart cold resistance to plants.

PS Claim 2; Pages 47-54; 70pp; Japanese.

CC O94645 encodes R76711 a recombinant cold-resistant pyruvate phosphate

CC dikinase (PPDK). The recombinant PPK can be used in the prodn. of

CC transformed cold-resistant plants, e.g. maize, deccan grass, millet

SQ Sequence 955 AA;

Query Match 43.8%; Score 63; DB 14; Length 955;

Best Local Similarity 46.7%; Pred. No. 2.45e+01;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 645 ehmfassderikavr 659

QY 4 DEKFYLSNGRIQAVR 18

# RESULT 8

ID W56265 standard; Protein; 350 AA.

AC W56265; 350 AA (first entry)

DE Mouse recombinase muREC2 protein.

KW Recombinase; REC2; hREC2; homologous recombination; irradiation;

OS Mus musculus.

PN W09811214-A1.

PD 19-MAR-1998.

PF 11-SEP-1997; IB1217.

PR 11-SEP-1996; US-025929.

PA (CORR ) CORNELL RES FOUND INC.

PA (UYJE-) UNIV JEFFERSON THOMAS.

PI Holloman WK, Kmiec EB, Rice MC, Shu Z, Smith ST;

DR WPI; 98-207378/18.

DR N-PSDB: V22704.

PT Mammalian and human REC2 genes encoding ATP-dependent recombinase(s)

PT - useful for the sensitisation of cells to the apoptotic effects of

PT irradiation

PS Disclosure; Fig 1E-F; 88pp; English.

CC This polypeptide comprises the murine homologue, muREC2, of the

CC ATP-dependent recombinase of Ustilago maydis (UsREC2). Its amino

CC acid sequence was deduced from a cDNA clone (see V22704) isolated

CC on the basis of homology to UmREC2. Overexpression of mammalian

CC REC2 (see also W56264) in a cell facilitates homologous

CC recombination between an exogenous DNA and the genome of a cell.

CC It also sensitises the cell to DNA damage, e.g. by gamma or UV

CC radiation or from cytotoxic agents. This causes the cell to

CC undergo apoptosis in response to DNA damage. The muREC2 gene can be

CC used in the construction of transgenic animals and knock-out mice

CC useful as models for tumorigenesis induced by chemical pathogens.

SQ Sequence 350 AA;

Query Match 42.4%; Score 61; DB 31; Length 350;

Best Local Similarity 33.3%; Pred. No. 3.92e+01;

Matches 6; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 164 fteeklltssrvhlc 181

QY 1 YNDEKFYLSNGRIQAVR 18

# RESULT 9

ID W79304 standard; Protein; 25 AA.

AC W79304;

DT 24-NOV-1998 (first entry)

DE Staphylococcus aureus GTP-binding protein.

KW GTP-binding protein; treatment; Synechocystis; prevention;

OS Staphylococcus aureus.

PN W09823738-A2.

PD 04-JUN-1998.

PF 24-NOV-1997; U22092.

PR 25-NOV-1996; US-031469.

PA (SMIK ) SMITHKLINE BEECHAM CORP.

PI Warren RL;

DR WPI; 98-322718/28.

DR N-PSDB: V59871.

PT New nucleic acid from Staphylococcus aureus NCIMB 40771 - useful



SQ Sequence 90 AA;

Query Match 40.3%; Score 58; DB 24; Length 90;  
 Best Local Similarity 22.2%; Pred. No. 7.85e+01;  
 Matches 4; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Db 24 fdgderfyvldkketi 41  
 : : ||| : : : :  
 QY 1 YNNDKFXLSNGRIQAVR 18

RESULT 13  
 ID R13117 standard; Protein; 258 AA.  
 AC R13117;  
 DT 08-OCT-1991 (first entry)  
 DE Cholera toxin A1 fragment.  
 KW Hybrid; fusion; membrane translocation; binding region; HIV;  
 KW infection; toxin; steroid; hormone; monoclonal antibody; antigen;  
 KW diphtheria; exotoxin; phenylketonuria; cholera; interleukin; IL-2;  
 KW protease; epidermal growth factor; ricin; tetanus; hexosaminidase;  
 KW Shiga-like toxin A; ligand; insulin; nuclease.  
 OS Vibrio cholera.  
 PN WO9109871-A.  
 PD 11-JUL-1991.  
 PF 21-DEC-1990; U07619.  
 PR 22-DEC-1989; US-456095.  
 PR 14-JUN-1990; US-538276.  
 PA (SERA-) SERAGEN INC.  
 PI Murphy JR;  
 DR WPI; 91-222845/30.  
 DR N-PSDB; Q12710.  
 PT Hybrid molecules for targeting chemical entity to cell - have  
 PT membrane translocating and cell binding-regions and used to  
 PT treat HIV infection, genetic enzyme-deficiency disorders etc.  
 PS Disclosure; Fig 5(1-2); 59pp; English.  
 CC Hybrid molecules are produced by covalently linking  
 CC (1) a portion (A) of the binding domain of a cell-binding ligand,  
 CC allowing binding of the mol. to an animal cell;  
 CC (2) a portion (B) of a translocation domain of a protein able to  
 CC translocate (C) across the cell cytoplasmic membrane, and  
 CC and (3) a portion (C) which is to be introduced into the cell.  
 CC (A) is derived from a steroid or polypeptide hormone, a single-chain  
 CC analogue of a monoclonal antibody able to bind an antigen expressed  
 CC on the cell surface, or a polypeptide toxin.  
 CC (B) is derived from a toxin (e.g. diphtheria toxin or Pseudomonas  
 CC exotoxin A).  
 CC (A) may be derived from insulin, interleukins 2, 3 or 6 or  
 CC epidermal growth factor.  
 CC Suitable enzymes in (C) include cholera toxin (represented in this  
 CC sequence), ricin, tetanus toxin, hexosaminidase A, protease,  
 CC nuclease, etc.  
 CC Specified examples are CT-A/DT-B'/IL-2, SLTA/DT-B'/IL-2,  
 CC ricin A/DT-B'/IL-2, HIVP-BP/DT-B'/IL-2 and the phenylalanine  
 CC hydroxylase-DT-B' or their biologically active mutants.  
 CC (CT-A= cholera toxin, DT-B'= truncated diphtheria toxin,  
 CC SLTA= Shiga-like toxin A; HIVP-BP= HIV protease binding protein.  
 CC See also Q12710-12.  
 SQ Sequence 258 AA;

Query Match 40.3%; Score 58; DB 3; Length 258;  
 Best Local Similarity 46.2%; Pred. No. 7.85e+01;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 17 yanddklyradsr 29  
 : : ||| : : : :  
 QY 1 YNNDKFXLSNGR 13

RESULT 14  
 ID W25785 standard; Protein; 258 AA.  
 AC W25785;  
 DT 27-MAR-1998 (first entry)  
 DE Vibrio cholerae toxin.

Cholera toxin A1; hybrid protein; cell delivery;  
 cell binding ligand; translocation domain; diphtheria toxin B';  
 interleukin-2; leukaemia; therapy.  
 OS Vibrio cholerae.  
 FH Key  
 FT Location/Qualifiers  
 Misc\_difference 120  
 /note= "encoded by GGG"  
 FT  
 Misc\_difference 206  
 /note= "encoded by GGG"  
 FT  
 US5668255-A.  
 PN  
 PD 16-SEP-1997.  
 PD 618199.  
 PR 07-JUN-1984; US-722484.  
 PR 27-JUN-1991; US-722484.  
 PR 07-JUN-1984; US-618199.  
 PR 25-APR-1985; US-726808.  
 PR 07-JUN-1985; US-742554.  
 PR 22-DEC-1989; US-456095.  
 PR 14-JUN-1990; US-538276.  
 PR 04-AUG-1993; US-102387.  
 PA (SERA-) SERAGEN INC.  
 PI Murphy JR;  
 DR WPI; 97-470103/43.  
 DR N-PSDB; T91636.  
 PT New hybrid molecules for delivery of agents to cells - comprise a  
 PT binding domain of a cell binding ligand and a portion of a  
 PT translocation domain of a protein.  
 PS Example 2; Fig 5A-B; 30pp; English.  
 CC This polypeptide comprises the cholera toxin encoded by Vibrio  
 CC cholerae. DNA (see T91636) encoding an enzymatically active  
 CC A1 fragment of the cholera toxin was used to construct a cholera  
 CC toxin A1-diphtheria toxin B'-interleukin-2 (CTA-DTB-IL2) gene that  
 CC was expressed in E. coli. The hybrid protein can be isolated and  
 CC used e.g. as an adjunct to treatment with DTB-IL2 hybrid, which  
 CC targets the cell-killing ability of diphtheria toxin to cells  
 CC bearing the IL2 receptor, such as certain leukaemic T cells.  
 CC Prior application of the CTA-DTB-IL2 hybrid alleviates the  
 CC problem of endogenous IL2 competing with DTB-IL2 hybrid for IL2  
 CC receptors on T cells. Claimed hybrid proteins comprise a  
 CC translocation domain and a cell binding domain from e.g. a hormone,  
 CC growth factor or polypeptide toxin. The hybrid molecules can be  
 CC used for the delivery of agents (e.g. therapeutic genes, toxins,  
 CC detectable labels) into cells. The use of a translocation  
 CC mechanism ensures that the hybrid will be effective in relatively  
 CC low doses, since a high proportion of the substance of interest  
 CC will be taken into the targeted cells. The hybrid molecules can be  
 CC manufactured as a single hybrid recombinant protein, permitting  
 CC reproducibility, consistency, and the precise control of  
 CC composition.  
 SQ Sequence 258 AA;

Query Match 40.3%; Score 58; DB 26; Length 258;  
 Best Local Similarity 46.2%; Pred. No. 7.85e+01;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 17 yanddklyradsr 29  
 : : ||| : : : :  
 QY 1 YNNDKFXLSNGR 13

RESULT 15  
 ID R96126 standard; Protein; 1169 AA.  
 AC R96126;  
 DT 28-OCT-1996 (first entry)  
 DE Bacillus thuringiensis crystal protein.  
 KW Bacillus thuringiensis; crystal protein; delta toxin; insecticide;  
 KW Lepidoptera; Coleoptera; crop protection; rice; wheat; beans; tea;  
 KW sugarcane; cauliflower; cabbage; apple; citrus fruit.  
 OS Bacillus thuringiensis var. japonensis, strain N141.  
 PN EP-7111834-A2.  
 PD 15-MAY-1996.  
 PF 13-OCT-1995; 307293.  
 PR 14-OCT-1994; JP-276082.  
 PA (NISC) NISSAN CHEM IND LTD.

PI Arai S, Iizuka T, Miyake T, Niizeki M, Tagawa M;  
DR WPI: 96-232099/24.  
PT N-PSDB: T27148.  
PT B. thuringiensis var japonensis strain N141 insecticidal crystal  
protein - used to protect plant from damage by pest, partic,  
lepidoptera or coleoptera insects  
PS Claim 3; Page 12-16; 20pp; English.  
CC The crystal protein is toxic to insects of the order Lepidoptera or  
CC Coleoptera, it is therefore useful in the protection of crops which  
CC are subject to infestation with lepidopterous or coleopterous insect  
CC pests. The crystal protein is formulated into a spray so that the  
CC protein agent is applied to crops at an amount of 0.1 to 5kg per  
CC hectare. Plants to be protected by such methods include  
CC vegetables such as cauliflower and cabbage, fruit trees such as  
CC citrus and apples, grains such as rice, wheat and beans and  
CC industrial crops such as tea and sugarcane.  
SQ Sequence 1169 AA;

Query Match 40.3%; Score 58; DB 18; Length 1169;  
Best Local Similarity 42.9%; Pred. No. 7.85e+01;  
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 435 dtnfyllsgqvsqi 448

QY |.:|||||:|:|:|:  
4 DEKFYLSNGRIQAV 17

Search completed: Mon Aug 23 13:04:37 1999  
Job time : 33 secs.



```

SUMMARY          immunoglobulin homology
                  #length 62 #checksum 7560

Query Match      61.1%; Score 88; DB 2; Length 62;
Best Local Similarity 47.4%; Pred. No. 6.83e-05;
Matches          9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 3
ENTRY   D34513      #type fragment
TITLE   MHC class II histocompatibility antigen DQ alpha chain
         (PDQA4.Y) - chimpanzee (fragment)
ORGANISM Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
DATE     22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
         08-Nov-1996
ACCESSIONS D34513; 136989
REFERENCE   A34512
          Gyllenstein, U.B.; Erlich, H.A.
          Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
          #journal
          #title Ancient roots for polymorphism at the HLA-DQalpha locus in
          primates.
#cross-references MUID:90099387
#accession D34513
#status     nucleic acid sequence not shown; not compared with
            conceptual translation

#molecule_type DNA
#residues   1-62 #label GYL

REFERENCE   I36899
          Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.;
          Jonker, M.; Bontrop, R.E.
          Immunogenetics (1992) 36:71-78
          #journal
          #title Evolutionary relationships among the primate MHC-DQ1 and
          DQ2 alleles.
#cross-references MUID:92307745
#accession I36969
#status     preliminary; translated from GB/EMBL/DBJ

#molecule_type DNA
#residues   1-62 #label RES
#cross-references GB:M76191; NID:q176844; PID:q176845
CLASSIFICATION #superfamily class II histocompatibility antigen;
immunoglobulin homology
#length 62 #checksum 7510

SUMMARY
Query Match      56.3%; Score 81; DB 2; Length 62;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches          8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 4
ENTRY   H34512      #type fragment
TITLE   MHC class II histocompatibility antigen HLA-DQ alpha 1 chain
         (HDOA4.3) - human (fragment)
ORGANISM Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
DATE     22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
         29-Aug-1997
ACCESSIONS H34512
REFERENCE   A34512
          Gyllenstein, U.B.; Erlich, H.A.
          Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
          #journal
          #title Ancient roots for polymorphism at the HLA-DQalpha locus in
          primates.
#cross-references MUID:90099387
#accession H34512
#status     nucleic acid sequence not shown; not compared with
            conceptual translation

SUMMARY          immunoglobulin homology
                  #length 62 #checksum 7560

Query Match      61.1%; Score 88; DB 2; Length 62;
Best Local Similarity 47.4%; Pred. No. 6.83e-05;
Matches          9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 3
ENTRY   D34513      #type fragment
TITLE   MHC class II histocompatibility antigen DQ alpha chain
         (PDQA4.Y) - chimpanzee (fragment)
ORGANISM Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
DATE     22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
         08-Nov-1996
ACCESSIONS D34513; 136989
REFERENCE   A34512
          Gyllenstein, U.B.; Erlich, H.A.
          Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
          #journal
          #title Ancient roots for polymorphism at the HLA-DQalpha locus in
          primates.
#cross-references MUID:90099387
#accession D34513
#status     nucleic acid sequence not shown; not compared with
            conceptual translation

#molecule_type DNA
#residues   1-62 #label GYL

REFERENCE   I36899
          Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.;
          Jonker, M.; Bontrop, R.E.
          Immunogenetics (1992) 36:71-78
          #journal
          #title Evolutionary relationships among the primate MHC-DQ1 and
          DQ2 alleles.
#cross-references MUID:92307745
#accession I36969
#status     preliminary; translated from GB/EMBL/DBJ

#molecule_type DNA
#residues   1-62 #label RES
#cross-references GB:M76191; NID:q176844; PID:q176845
CLASSIFICATION #superfamily class II histocompatibility antigen;
immunoglobulin homology
#length 62 #checksum 7510

SUMMARY
Query Match      56.3%; Score 81; DB 2; Length 62;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches          8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 4
ENTRY   H34512      #type fragment
TITLE   MHC class II histocompatibility antigen HLA-DQ alpha 1 chain
         (HDOA4.3) - human (fragment)
ORGANISM Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
DATE     22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
         29-Aug-1997
ACCESSIONS H34512
REFERENCE   A34512
          Gyllenstein, U.B.; Erlich, H.A.
          Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
          #journal
          #title Ancient roots for polymorphism at the HLA-DQalpha locus in
          primates.
#cross-references MUID:90099387
#accession H34512
#status     nucleic acid sequence not shown; not compared with
            conceptual translation

#molecule_type DNA
#residues   1-62 #label GYL

GENETICS
#gene      GDB:HLA-DQA1
#cross-references GDB:120638; OMIM:146880
#map_position 6p21.3-6p21.3
CLASSIFICATION #superfamily class II histocompatibility antigen;
immunoglobulin homology
#length 62 #checksum 7348

SUMMARY
Query Match      56.3%; Score 81; DB 2; Length 62;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches          8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 5
ENTRY   B35589      #type fragment
TITLE   MHC class II histocompatibility antigen HLA-DQ alpha-4.3
         Chain (LUV) - human (fragment)
ORGANISM #formal_name Homo sapiens #common_name man
DATE     14-Sep-1990 #sequence_revision 14-Sep-1990 #text_change
         12-Jul-1996
ACCESSIONS B35589
REFERENCE   A35589
          Horn, G.T.; Bugawan, T.L.; Long, C.M.; Erlich, H.A.
          Proc. Natl. Acad. Sci. U.S.A. (1988) 85:6012-6016
          #journal
          #title Allelic sequence variation of the HLA-DQ loci: relationship
          to serology and to insulin-dependent diabetes
          susceptibility.
#cross-references MUID:88320372
#accession B35589
#status     preliminary; nucleic acid sequence not shown; not
            compared with conceptual translation

#molecule_type DNA
#residues   1-63 #label HOR
CLASSIFICATION #superfamily class II histocompatibility antigen;
immunoglobulin homology
#length 63 #checksum 7816

SUMMARY
Query Match      56.3%; Score 81; DB 2; Length 63;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches          8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30
  :: |||||: || :||:|
QY 1 YNNDEKFYLSNGRIQAVRC 19

RESULT 6
ENTRY   HLHU3C      #type complete
TITLE   MHC class II histocompatibility antigen HLA-DQ alpha 1 chain
         precursor (allele DQA1*0501) - human
ALTERNATE_NAMES MHC class II histocompatibility antigen HLA-DQ alpha-4.1
                 Chain
ORGANISM #formal_name Homo sapiens #common_name man
DATE     17-May-1985 #sequence_revision 17-May-1985 #text_change
         05-Sep-1997
ACCESSIONS A02215; D27618; D30575; A32527; F34512; I68550; I59626
REFERENCE   A02215
          Schenning, L.; Larhammar, D.; Bill, P.; Wiman, K.; Jonsson,
          A.K.; Rask, L.; Peterson, P.A.
          EMBO J. (1984) 3:447-452
          #journal
          #title Both alpha and beta chains of HLA-DC class II
          histocompatibility antigens display extensive polymorphism
          in their amino-terminal domains.
#cross-references MUID:84182509
#accession A02215
#molecule_type mRNA
#residues   1-254 #label SCH
```

```

#accession      I59626
#status         translated from GB/EMBL/DBDJ
#molecule_type DNA
#residues       35-74 #label RE2
#cross-references EMBL:U03675; NID:g606783; PID:g606784
#note          this allele is designated HLA-DQA1*0502

GENETICS
#gene          GDB:HLA-DQA1
#cross-references GDB:120638; OMIM:146880
#map_position 6p21.3-6p21.3
CLASSIFICATION #superfamily class II histocompatibility antigen;
               immunoglobulin homology
               glycoprotein; heterodimer; transmembrane protein
KEYWORDS
FEATURE
1-23          #domain signal sequence #status predicted #label SIG\
24-254        #product MHC class II histocompatibility antigen HLA-DQ\
               alpha 1 chain #status predicted #label MAT\
24-216        #domain extracellular #status predicted #label EXT\
24-119        #domain alpha-1 #label EX1\
125-190       #domain immunoglobulin homology #label IMM\
217-239       #domain transmembrane #status predicted #label TM\
240-254       #domain intracellular #status predicted #label INT\
103,143       #binding_site carbohydrate (asn) (covalent) #status
               predicted\
132-188       #disulfide_bonds #status predicted
SUMMARY
#length 254 #molecular_weight 27805 #checksum 4535

Query Match      56.3%; Score 81; DB 1; Length 254;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 52 FDGEQFYVDLGRKETVVC 70
   : : : : :
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 7
ENTRY  MHC class II histocompatibility antigen HLA-DQ alpha 1 (DQw4
TITLE  specificity) precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change
      08-Sep-1997
ACCESSIONS A37044; A27583; G34512; C35589
REFERENCE  A37044
AUTHORS Jonsson, A.K.; Andersson, L.; Rask, L.
#journal Immunogenetics (1989) 30:232-234
#title Complete sequences of DQa1 and DQb1 cDNA clones corresponding
        to the DQw4 specificity.
#cross-references MUID:89379293
#accession A37044
#status preliminary
#molecule_type mRNA
#residues 1-254 #label JON
#cross-references GB:M33906
#note this allele is designated DQA1*0401
REFERENCE A27583
AUTHORS Hurley, C.K.; Gregeren, P.; Steiner, N.; Bell, J.; Hartzman,
        R.; Nepom, G.; Silver, J.; Johnson, A.H.
#journal J. Immunol. (1988) 140:885-892
#title Polymorphism of the HLA-D region in American blacks. A DR3
        haplotype generated by recombination.
#cross-references MUID:88116708
#accession A27583
#molecule_type mRNA
#residues 24-109 #label HUR
#cross-references GB:M18935; NID:g188260; PID:g188261
REFERENCE A34512
AUTHORS Gyllenstein, U.B.; Erlich, H.A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
#title Ancient roots for polymorphism at the HLA-DQA1 locus in
        primates.
#cross-references MUID:90099387

```

```

#accession G34512      nucleic acid sequence not shown; not compared with
#status      conceptual translation
#molecule_type DNA
#residues    41-102 #label GYL
#note        allele DQA4.2
REFERENCE
A35589
#authors     Horn, G.T.; Bugawan, T.L.; Long, C.M.; Erlich, H.A.
#journal     Proc. Natl. Acad. Sci. U.S.A. (1988) 85:6012-6016
#title       Allelic sequence variation of the HLA-DQ loci: relationship
             to serology and to insulin-dependent diabetes
             susceptibility.
#cross-references MUID:88320372
#accession   C35589
#status      preliminary; nucleic acid sequence not shown; not
             compared with conceptual translation
#molecule_type DNA
#residues    41-102 #label HOR
GENETICS
#gene        GDB:HLA-DQA1
#map_position 6p21.3-6p21.3
#classification #superfamily class II histocompatibility antigen;
             immunoglobulin homology
KEYWORDS
heterodimer
SUMMARY
#length 254 #molecular-weight 27836 #checksum 4214
Query Match      56.3%; Score 81; DB 2; Length 254;
Best Local Similarity 42.1%; Pred. No. 1.72e-03;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Db 52 FDGEQFYVDLGRKETVVC 70
:: ||||| ::|||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 8
ENTRY #
TITLE H71325 #type complete
#organism #spirochete
#formal_name Treponema pallidum subsp. pallidum #common_name
             syphilis spirochete
DATE 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change
17-Mar-1999
ACCESSIONS H71325
REFERENCE A71250
#authors     Sutton, G.G.; Dodson, R.; Gwinn, M.; Hickey, E.K.; Clayton,
             R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod,
             M.P.; Salzberg, S.; Peterson, J.; Khalak, H.; Richardson,
             D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDonald,
             L.; Artiaich, P.; Bowman, C.; Cotton, M.D.; Fujii, C.;
             Garland, S.; Hatch, B.; Horst, K.; Roberts, K.; Watthey,
             L.; Weidman, J.; Smith, H.O.; Venter, J.C.
#journal     Science (1998) 281:373-388
#title       Complete genome sequence of Treponema pallidum, the syphilis
             spirochete.
#cross-references MUID:98332770
#accession   H71325
#status      preliminary; nucleic acid sequence not shown;
             translation not shown
#molecule_type DNA
#residues    1-430 #label COL
#cross-references GB:AE001220; GB:AE000520; NID:g3322705; PID:g3322710
#experimental_source strain Nichols
GENETICS TP0427
#gene        #superfamily H+-transporting ATP synthase alpha chain;
             H+-transporting ATP synthase alpha chain homology
#classification #length 430 #molecular-weight 47229 #checksum 6377
SUMMARY
Query Match      52.1%; Score 75; DB 2; Length 430;
Best Local Similarity 50.0%; Pred. No. 2.47e-02;

```

```

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Db 309 YTEGQFYLNKNGRIEP 324
| : ||| |||||
QY 1 YNDEKFYLSNGRIQA 16

RESULT 9
ENTRY #
TITLE I34513 #type fragment
#organism MHC class II histocompatibility antigen DQ alpha chain
             (DQA4) - gorilla (fragment)
#alternate_names MHC DQ-alpha 1
#formal_name Gorilla gorilla #common_name gorilla
DATE 22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
08-Nov-1996
ACCESSIONS I34513; I37065
REFERENCE A34512
#authors     Gyllenstein, U.B.; Erlich, H.A.
#journal     Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
#title       Ancient roots for polymorphism at the HLA-DQalpha locus in
             primates.
#cross-references MUID:90099387
#accession   I34513
#status      nucleic acid sequence not shown; not compared with
             conceptual translation
#molecule_type DNA
#residues    1-62 #label GYL
REFERENCE I36899
#authors     Kenter, M.; Otting, N.; Anholts, J.; Leunissen, J.A.M.;
             Jonker, M.; Bontrop, R.E.
#journal     Immunogenetics (1992) 36:71-78
#title       Evolutionary relationships among the primate MHC-DQa1 and
             DQA2 alleles.
#cross-references MUID:92307745
#accession   I37065
#status      preliminary; translated from GB/EMBL/DDJB
#molecule_type DNA
#residues    1-62 #label RES
#cross-references GB:M76198; NID:gi77071; PID:gi77072
#classification #superfamily class II histocompatibility antigen;
             immunoglobulin homology
SUMMARY #length 62 #checksum 7464
Query Match      50.7%; Score 73; DB 2; Length 62;
Best Local Similarity 36.8%; Pred. No. 5.85e-02;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
Db 12 FDGEQFYVDLGRKETVVC 30
:: ||||| ::|||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 10
ENTRY #
TITLE C70454 #type complete
#organism GTP-binding protein HflX. Aquifex aeolicus
#formal_name Aquifex aeolicus
DATE 08-May-1998 #sequence_revision 08-May-1998 #text_change
21-Dec-1998
ACCESSIONS C70454
REFERENCE A70300
#authors     Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;
             Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.;
             Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short,
             J.M.; Olson, G.J.; Swanson, R.V.
#journal     Nature (1998) 392:353-358
#title       The complete genome of the hyperthermophilic bacterium
             Aquifex aeolicus.
#cross-references MUID:98196666
#accession   C70454
#status      preliminary; nucleic acid sequence not shown;
             translation not shown
#molecule_type DNA
#residues    1-370 #label AQF

```



```

##cross-references GB:AE000762; NID:g2984163; PID:g2984170; GB:AE000657
##experimental_source strain VF5
GENETICS
#gene hflx
CLASSIFICATION #superfamily GTP-binding protein hflx; translation elongation
                 factor Tu homology
KEYWORDS      GTP binding; P-loop
FEATURE
197-319      #domain translation elongation factor Tu homology #label
              EFL
203-210      #region nucleotide-binding motif A (P-loop)\
316-319      #region GTP-binding NKXD motif
SUMMARY      #length 370 #molecular-weight 42100 #checksum 3162
Query Match  47.9%; Score 69; DB 2; Length 370;
Best Local Similarity 40.0%; Pred. No. 3.17e-01;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
Db 48 DQYIIGGQRAQEI 62
   |||: ||: ||: ||
Qy 4 DEKYLNGRIQAVR 18

RESULT 11
ENTRY      T02171      #type complete
TITLE      hypothetical protein Lu-ECAM-1 - bovine
ORGANISM   #formal_name Bos primigenius taurus #common_name cattle
DATE       05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02171
REFERENCE   Z14590
#authors   Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
              endothelial chloride channel.
#accession T02171
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-794 #label ELB
##cross-references EMBL:AF001262; NID:g2623764; PID:g2623765
##experimental_source lung
SUMMARY     #length 794 #molecular-weight 88509 #checksum 6853

Query Match  47.9%; Score 69; DB 2; Length 794;
Best Local Similarity 50.0%; Pred. No. 3.17e-01;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
Db 168 YNVDQPFYISRKNTIATRC 187
   |||: ||: ||: ||
Qy 1 YNNDKPYLS-NGRIQAVRC 19

RESULT 12
ENTRY      T02152      #type complete
TITLE      Lu-ECAM-1 protein - bovine
ORGANISM   #formal_name Bos primigenius taurus #common_name cattle
DATE       05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02152
REFERENCE   Z14590
#authors   Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
              endothelial chloride channel.
#accession T02152
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-820 #label ELB
##cross-references EMBL:AF001263; NID:g2623766; PID:g2623767
##experimental_source lung
SUMMARY     #length 820 #molecular-weight 91464 #checksum 3104

```

```

Query Match  47.9%; Score 69; DB 2; Length 820;
Best Local Similarity 50.0%; Pred. No. 3.17e-01;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
Db 168 YNVDQPFYISRKNTIATRC 187
   |||: ||: ||: ||
Qy 1 YNNDKPYLS-NGRIQAVRC 19

RESULT 13
ENTRY      T02205      #type complete
TITLE      Lu-ECAM-1 protein - bovine
ORGANISM   #formal_name Bos primigenius taurus #common_name cattle
DATE       05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change
ACCESSIONS T02205
REFERENCE   Z14590
#authors   Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission submitted to the EMBL Data Library, April 1997
#description Cloning and characterization of Lu-ECAM-1 suggest it is an
              endothelial chloride channel.
#accession T02205
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-905 #label ELB
##cross-references EMBL:AF001261; NID:g2623762; PID:g2623763
##experimental_source lung
SUMMARY     #length 905 #molecular-weight 101005 #checksum 5557

Query Match  47.9%; Score 69; DB 2; Length 905;
Best Local Similarity 50.0%; Pred. No. 3.17e-01;
Matches 10; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
Db 168 YNVDQPFYISRKNTIATRC 187
   |||: ||: ||: ||
Qy 1 YNNDKPYLS-NGRIQAVRC 19

RESULT 14
ENTRY      E34514      #type fragment
TITLE      MHC class II histocompatibility antigen DQ alpha chain 3 -
              rhesus macaque (fragment)
ORGANISM   #formal_name Macaca mulatta #common_name rhesus macaque
DATE       22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
ACCESSIONS E34514
REFERENCE   A34512
#authors   Gyllenstein, U.B.; Erlich, H.A.
#journal   Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
#title     Ancient roots for polymorphism at the HLA-DQAalpha locus in
              primates.
#cross-references MUID:90099387
#accession E34514
#status    nucleic acid sequence not shown; not compared with
              conceptual translation
##molecule_type DNA
##residues  1-62 #label GYL
CLASSIFICATION #superfamily class II histocompatibility antigen;
              immunoglobulin homology
SUMMARY     #length 62 #checksum 8894

Query Match  47.2%; Score 68; DB 2; Length 62;
Best Local Similarity 38.9%; Pred. No. 4.79e-01;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Db 12 FDGDEQFYVDLGRKTVW 29
   |||: ||: ||: ||
Qy 1 YNNDKPYLSNGRIQAVR 18

RESULT 15

```

```

ENTRY      F34514      #type fragment
TITLE      MHC class II histocompatibility antigen DQ alpha chain 1 -
            hanuman langur (fragment)
ORGANISM    #formal_name Presbytis entellus #common_name hanuman langur
DATE        22-Jun-1990 #sequence_revision 26-Apr-1996 #text_change
            10-May-1996
ACCESSIONS  F34514
REFERENCE   A34512
            #authors Gyllenstein, U.B.; Erlich, H.A.
            #journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:9986-9990
            #title Ancient roots for polymorphism at the HLA-DQalpha locus in
            primates.
#cross-references MUID:90099387
#accession F34514
#status nucleic acid sequence not shown: not compared with
            conceptual translation
#molecule_type DNA
#residues 1-62 #label GYL
CLASSIFICATION #superfamily class II histocompatibility antigen;
immunoglobulin homology
SUMMARY      #length 62 #checksum 6700

Query Match 47.2%; Score 68; DB 2; Length 62;
Best Local Similarity 38.9%; Pred.No. 4.79e-01;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVW 29
   ::|||::||::|:
QY 1 YNNDERFYLSNGRIQAVR 18

```

Search completed: Mon Aug 23 13:03:44 1999  
Job time : 14 secs.

\*\*\*\*\*  
Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd  
\*\*\*\*\*  
MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:02:21 1999; MasPar time 3.41 Seconds  
Tabular output not generated.  
\*\*\*\*\*  
Title: >US-09-049-696-42  
Description: (1-19) from US09049696.pap  
Perfect Score: 144  
Sequence: 1 YNNDKFKYLSNGRIQAVRC 19  
Scoring table: PAM 150  
Gap 15  
Searched: 77977 seqs, 28268293 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: swiss:prot37  
1:swissprot  
Statistics: Mean 29.568; Variance 37.795; scale 0.782  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.  
\*\*\*\*\*  
SUMMARIES  
\*\*\*\*\*

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	81	56.3	254	1	HA27_HUMAN HLA CLASS II HISTOCOMP	1.59e-04
2	81	56.3	254	1	HA23_HUMAN HLA CLASS II HISTOCOMP	1.59e-04
3	70	48.6	903	1	ECLC_BOVIN EPITHELIAL CHLORIDE CH	3.73e-02
4	67	46.5	503	1	ALG2_YEAST GLYCOSYLTRANSFERASE AL	1.52e-01
5	64	44.4	408	1	Y889_SYNV3 HYPOTHETICAL 46.0 KD P	5.96e-01
6	63	43.8	947	1	PODK_MAIZE PYRUVATE PHOSPHATE DIK	9.30e-01
7	63	43.8	953	1	PODK_FLABI PYRUVATE PHOSPHATE DIK	9.30e-01
8	63	43.8	953	1	PODK_FLATR PYRUVATE PHOSPHATE DIK	9.30e-01
9	63	43.8	956	1	PODK_FLABR PYRUVATE PHOSPHATE DIK	9.30e-01
10	63	43.8	955	1	PODK_FLAPR PYRUVATE PHOSPHATE DIK	9.30e-01
11	62	43.1	255	1	HA2D_PIG SLA CLASS II HISTOCOMP	1.44e+00
12	62	43.1	505	1	PK76_SYNV3 PROBABLE SERINE/THREON	1.44e+00
13	62	43.1	949	1	PODK_MESCR PYRUVATE PHOSPHATE DIK	1.44e+00
14	61	42.4	350	1	R51B_MOUSE DNA REPAIR PROTEIN RAD	2.23e+00
15	61	42.4	1201	1	MFQ_MYXXA TRANSCRIPTION-REPAIR C	2.23e+00
16	60	41.7	634	1	YCX3_EUGGR HYPOTHETICAL 78.0 KD P	3.43e+00
17	59	41.0	232	1	HA24_HUMAN HLA CLASS II HISTOCOMP	5.23e+00
18	59	41.0	254	1	HA22_HUMAN HLA CLASS II HISTOCOMP	5.23e+00
19	59	41.0	255	1	HA21_HUMAN HLA CLASS II HISTOCOMP	5.23e+00
20	59	41.0	383	1	HIPO_CAMJE HIPURATE HYDROLASE (E	5.23e+00
21	59	41.0	473	1	L14B_LITER LEC14B PROTEIN.	5.23e+00
22	59	41.0	603	1	MTM1_HUMAN MYOTUBULARIN.	5.23e+00
23	59	41.0	4568	1	DYHC_CAEEL DYNEIN HEAVY CHAIN, CY	5.23e+00

RESULT ID	HA27_HUMAN	STANDARD	PRT	254 AA
AC	P05536;			
DT	01-NOV-1988 (REL. 09, CREATED)			
DT	01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)			
DT	01-JAN-1990 (REL. 13, LAST ANNOTATION UPDATE)			
DE	HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(W3) ALPHA CHAIN PRECURSOR.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
RC	PRIMATES; CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE: 87224159.			
RA	SCHIFFENBAUER J., DIDIER D.K., KLEARMAN M., RICE K., SHUMAN S.,			
RA	TIEBER V.L., KITTLESEN D.J., SCHWARTZ B.D.;			
RT	"Complete sequence of the HLA DQ alpha and DQ beta cDNA from a			
RT	DR5/DQw3 cell line."			
RL	J. IMMUNOL. 139:228-233(1987).			
CC	.....			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	.....			
DR	EMBL: M16995; G307252;			
DR	PIR: A32527; A32527.			
DR	PROSITE: PS00290; IG_MHC; 1.			
DR	PFAM: PF00847; ig; 1.			
DR	PFAM: PF00993; MHC_II_alpha; 1.			
DR	HSSP: P01910; 1IAK.			
KW	MHC II; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.			
FT	SIGNAL 1 23			
FT	CHAIN 24 254			
FT	DOMAIN 24 119			
FT	DOMAIN 120 203			
FT	DOMAIN 204 216			
FT	TRANSMEM 217 239			
FT	DOMAIN 240 254			
FT	DISULFID 132 188			
FT	CARBOHYD 103 103			
FT	CARBOHYD 143 143			
SQ	SEQUENCE 254 AA; 27845 MW; 3CA08E4D CRC32;			

Query Match 56.3%; Score 81; DB 1; Length 254;  
Best Local Similarity 42.1%; Pred. No. 1.59e-04;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Dd 52 FDGDEQFYVDLGRKETVWC 70  
:: || ||: ||:||  
QY 1 YNNDKFKYLSNGRIQAVRC 19

RESULT 2  
ID HA23\_HUMAN STANDARD; PRT; 254 AA.  
AC F01909;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(3) ALPHA CHAIN PRECURSOR (DC-ALPHA)  
DE DE  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A. (CLONE PII-ALPHA-5).  
RX MEDLINE: 84182509.  
RA SCHENNING L., LAHAMMAR D., BILL P., WIMAN K., JONSSON A.-K., RASK L., PETERSON P.A.;  
RT "Both alpha and beta chains of HLA-DC class II histocompatibility antigens display extensive polymorphism in their amino-terminal domains.";  
RT EMBO J. 3:447-452(1984).  
RL CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: K01160; G307243; .  
CC EMBL: X00370; . NOT\_ANNOTATED\_CDS.  
CC PIR: A02215; HLHU3C.  
CC DR PROSITE: PS00290; IG\_MHC; 1.  
CC DR PFAM: PF00047; Ig; 1.  
CC DR PFAM: PF00993; MHC\_II\_alpha; 1.  
CC DR HSP: P01910; IIAK.  
CC KW MHC II; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.  
CC FT SIGNAL 1 23  
CC FT CHAIN 24 254  
CC FT CD(3) ALPHA CHAIN.  
CC FT DOMAIN 24 119  
CC FT DOMAIN 120 203  
CC FT DOMAIN 204 216  
CC FT TRANSMEM 217 239  
CC FT DOMAIN 240 254  
CC FT DISULFID 132 188  
CC FT CARBOHYD 103 103  
CC FT CARBOHYD 143 143  
CC FT CARBOHYD 143 143  
SQ SEQUENCE 254 AA; 27805 MW; 05C2CC5A CRC32;

Query Match 56.3%; Score 81; DB 1; Length 254;  
Best Local Similarity 42.1%; Pred. No. 1.59e-04;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Dd 52 FDGDEQFYVDLGRKETVWC 70  
:: || ||: ||:||  
QY 1 YNNDKFKYLSNGRIQAVRC 19

RESULT 3  
ID ECIC\_BOVIN STANDARD; PRT; 903 AA.  
AC P54281;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)



```

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U08400; G606809; -
CC PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; 1.
CC PROSITE; PS00742; PEP_ENZYMES_2; 1.
CC PFAM; PF00391; PEP-utilizers; 1.
CC PFAM; PF01326; PDPK_N_term; 1.
CC HSP; P22983; 1DKK.
CC TRANSFERASE; KINASE; ATP-BINDING; PHOSPHORYLATION; PHOTOSYNTHESIS;
KW CHLOROPLAST; TRANSIT PEPTIDE.
KW TRANSIT 1 76 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 77 953 PYRUVATE,PHOSPHATE DIKINASE.
FT MOD_RES 533 533 PHOSPHORYLATION (LIGHT-DEPENDENT) (BY
FT SIMILARITY).
FT MOD_RES 535 535 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 953 AA; 104002 MW; 6949BC1D CRC32;
Query Match 43.8%; Score 63; DB 1; Length 953;
Best Local Similarity 46.7%; Pred. No. 9.30e-01;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Db 643 EHMFFASDERIKAVR 657
: : : : :
QY 4 DEKFLYSLNGRIQAVR 18
-----
RESULT 8
ID PODK_FLATR STANDARD; PRT; 953 AA.
AC P22221;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DT DE (PYRUVATE,PHOSPHATE DIKINASE PRECURSOR (EC 2.7.9.1)
GN (PYRUVATE, ORTHOPHOSPHATE DIKINASE).
OS PPDK.
GN FLAVERIA TRINERVIA.
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; ASTERALES; ASTERACEAE; FLAVERIA.
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=LEAF.
RX MEDLINE; 91032158.
RA ROSCHE E., WESTHOFF P.;
RT "Primary structure of pyruvate, orthophosphate dikinase in the
RT dicotyledonous C4 plant Flaveria trinervia.";
RL FEBS LETT. 273:116-121(1990).
CC -!- FUNCTION: FORMATION OF PHOSPHOENOLPYRUVATE, WHICH IS THE
CC PRIMARY ACCEPTOR OF CO2 IN C4 AND SOME CRASSULACEAN ACID
CC METABOLISM PLANTS.
CC -!- CATALYTIC ACTIVITY: ATP + PYRUVATE + PHOSPHATE = AMP +
CC PHOSPHOENOLPYRUVATE + DIPHOSPHATE.
CC -!- ENZYME REGULATION: CONTROLLED BY LIGHT THROUGH A PHOSPHORYLATION/
CC DEPHOSPHORYLATION MECHANISM.
CC -!- PATHWAY: C4 PHOTOSYNTHETIC PATHWAY.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- TISSUE SPECIFICITY: IN MESOPHYLL CELLS OF LEAVES.
CC -!- SIMILARITY: TO OTHER PEP-UTILIZING ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57141; G18462; -
CC PIR; S12894; S12894.

```

DR PROSITE; PS00370; PEP\_ENZYMES\_PHOS\_SITE; 1.  
 DR PROSITE; PS00742; PEP\_ENZYMES\_2; 1.  
 DR PFAM; PF00391; PEP-utilizers; 1.  
 DR PFAM; PF01326; PPK\_N-term; 1.  
 DR HSP; P22983; IDIK.  
 KW TRANSFERASE; KINASE; ATP-BINDING; PHOSPHORYLATION; PHOTOSYNTHESIS;  
 FW CHLOROPLAST; TRANSIT PEPTIDE.  
 FT TRANSIT 1 76 CHLOROPLAST.  
 FT CHAIN 77 953 PYRUVATE PHOSPHATE DIKINASE.  
 FT MOD\_RES 533 533 PHOSPHORYLATION (LIGHT-DEPENDENT) (BY  
 FT SIMILARITY).  
 FT MOD\_RES 535 535 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 953 AA; 103949 MW; BA1B00C9 CRC32;  
 Query Match 43.8%; Score 63; DB 1; Length 953;  
 Best Local Similarity 46.7%; Pred.No. 9.30e-01;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 643 EHMFFASDERIKAVR 657  
 : : : : :  
 QY 4 DERFYLNGRIQAVR 18

RESULT 9  
 ID PODK\_FLAPR STANDARD; PRT; 955 AA.  
 AC Q39734;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PYRUVATE,PHOSPHATE DIKINASE PRECURSOR (EC 2.7.9.1)  
 DE (PYRUVATE,ORTHOPHOSPHATE DIKINASE) (COLD STABLE PYRUVATE,  
 DE ORTHOPHOSPHATE DIKINASE).  
 GN PPK.  
 OS FLAVERIA BROWNII.  
 OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
 OC ASTERIDAE; ASTERALES; ASTERACEAE; FLAVERIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE; 95284353.  
 RA USAMI S., OHTA S., KOMARI T., BURNELL J.N.;  
 RT "Cold stability of pyruvate, orthophosphate dikinase of Flaveria  
 RT brownii.";  
 RL PLANT MOL. BIOL. 27:969-980(1995).  
 CC -!- FUNCTION: FORMATION OF PHOSPHOENOLPYRUVATE, WHICH IS THE  
 CC PRIMARY ACCEPTOR OF CO2 IN C4 AND SOME CRASSULACEAN ACID  
 CC METABOLISM PLANTS.  
 CC -!- CATALYTIC ACTIVITY: ATP + PYRUVATE + PHOSPHATE = AMP +  
 CC PHOSPHOENOLPYRUVATE + DIPHOSPHATE.  
 CC -!- ENZYME REGULATION: CONTROLLED BY LIGHT THROUGH A PHOSPHORYLATION/  
 CC DEPHOSPHORYLATION MECHANISM (BY SIMILARITY).  
 CC -!- PATHWAY: C4 PHOTOSYNTHETIC PATHWAY.  
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- SIMILARITY: TO OTHER PEP-UTILIZING ENZYMES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U08399; G606807;  
 DR PROSITE; PS00370; PEP\_ENZYMES\_PHOS\_SITE; 1.  
 DR PROSITE; PS00742; PEP\_ENZYMES\_2; 1.  
 DR PFAM; PF00391; PEP-utilizers; 1.  
 DR PFAM; PF01326; PPK\_N-term; 1.  
 DR HSP; P22983; IDIK.  
 KW TRANSFERASE; KINASE; ATP-BINDING; PHOSPHORYLATION; PHOTOSYNTHESIS;  
 FW CHLOROPLAST; TRANSIT PEPTIDE.

FT TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 79 955 PYRUVATE,PHOSPHATE DIKINASE.  
 FT MOD\_RES 535 535 PHOSPHORYLATION (LIGHT-DEPENDENT) (BY  
 FT SIMILARITY).  
 FT MOD\_RES 537 537 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 955 AA; 104224 MW; ABC066F2 CRC32;  
 Query Match 43.8%; Score 63; DB 1; Length 955;  
 Best Local Similarity 46.7%; Pred.No. 9.30e-01;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 645 EHMFFASDERIKAVR 659  
 : : : : :  
 QY 4 DERFYLNGRIQAVR 18

RESULT 10  
 ID PODK\_FLAPR STANDARD; PRT; 956 AA.  
 AC Q42736;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE PYRUVATE,PHOSPHATE DIKINASE PRECURSOR (EC 2.7.9.1)  
 DE (PYRUVATE,ORTHOPHOSPHATE DIKINASE).  
 GN PPK OR PDK.  
 OS FLAVERIA PRINGLEI.  
 OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;  
 OC ASTERIDAE; ASTERALES; ASTERACEAE; FLAVERIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RX MEDLINE; 95036056.  
 RA ROSCHE E., STREUBEL M., WESTHOFF P.;  
 RT "Primary structure of the photosynthetic pyruvate orthophosphate  
 RT dikinase of the C3 plant Flaveria pringlei and expression analysis of  
 RT pyruvate orthophosphate dikinase sequences in C3, C3-C4 and C4  
 RT Flaveria species";  
 RL PLANT MOL. BIOL. 26:763-769(1994).  
 CC -!- FUNCTION: FORMATION OF PHOSPHOENOLPYRUVATE.  
 CC -!- CATALYTIC ACTIVITY: ATP + PYRUVATE + PHOSPHATE = AMP +  
 CC PHOSPHOENOLPYRUVATE + DIPHOSPHATE.  
 CC -!- ENZYME REGULATION: CONTROLLED BY LIGHT THROUGH A PHOSPHORYLATION/  
 CC DEPHOSPHORYLATION MECHANISM (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- SIMILARITY: TO OTHER PEP-UTILIZING ENZYMES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X75516; G577776;  
 DR PROSITE; PS00370; PEP\_ENZYMES\_PHOS\_SITE; 1.  
 DR PROSITE; PS00742; PEP\_ENZYMES\_2; 1.  
 DR PFAM; PF00391; PEP-utilizers; 1.  
 DR PFAM; PF01326; PPK\_N-term; 1.  
 DR HSP; P22983; IDIK.  
 KW TRANSFERASE; KINASE; ATP-BINDING; PHOSPHORYLATION; PHOTOSYNTHESIS;  
 FW CHLOROPLAST; TRANSIT PEPTIDE.  
 FT TRANSIT 1 79 CHLOROPLAST (BY SIMILARITY).  
 FT CHAIN 80 956 PYRUVATE,PHOSPHATE DIKINASE.  
 FT MOD\_RES 536 536 PHOSPHORYLATION (LIGHT-DEPENDENT) (BY  
 FT SIMILARITY).  
 FT MOD\_RES 538 538 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 956 AA; 104554 MW; 793EB24B CRC32;  
 Query Match 43.8%; Score 63; DB 1; Length 956;  
 Best Local Similarity 46.7%; Pred.No. 9.30e-01;

```

Matches      7:  Conservative      4:  Mismatches      4:  Indels      0:  Gaps      0;

Db      646 EHMFFASDERIKAVR 660
QY      4 DEKFYLSNGRIQAVR 18

RESULT 11
ID HA2D_PIG STANDARD; PRT; 255 AA.
AC P15981;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE SLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ ALPHA CHAIN PRECURSOR
DE (HAPLOTYPE D).
OS SUS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE: 90129156.
RA HIRSCH F., SACHS D.H., GUSTAFSSON K., PRATT K., GERMANA S.,
RA LEGUERN C.;
RT "Class II genes of miniature swine. III. Characterization of an
RT expressed pig class II gene homologous to HLA-DQA.";
RL IMMUNOGENETICS 31:52-56(1990).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

-----
DR EMBL: M29939; G164565; -.
DR PROSITE: PS00290; IG_MHC; 1.
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00993; MHC_II_alpha; 1.
DR HSP: P01910; IIAK.
KW MHC II; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL.
FT SIGNAL 1 23
FT CHAIN 24 255
FT DOMAIN 24 120 SLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT DOMAIN 121 204 DQ ALPHA CHAIN (HAPLOTYPE D).
FT DOMAIN 205 217 EXTRACELLULAR ALPHA-1.
FT TRANSMEM 218 240 EXTRACELLULAR ALPHA-2.
FT DOMAIN 241 255 CONNECTING PEPTIDE.
FT DISULFID 133 189 CYTOPLASMIC.
FT CARBOHYD 104 104 BY SIMILARITY.
FT CARBOHYD 144 144 POTENTIAL.
SQ SEQUENCE 255 AA; 28067 MW; 60E3DC37 CRC32;

Query Match 43.1%; Score 62; DB 1; Length 255;
Best Local Similarity 33.3%; Pred. No. 1.44e+00;
Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db      52 FGDDEEYVDLGGKETVW 69
QY      1 YNNDKFYLSNGRIQAVR 18

RESULT 12
ID PK76_SVNY3 STANDARD; PRT; 505 AA.
AC P54735;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE SERINE/THREONINE-PROTEIN KINASE SLL0776 (BC 2.7.1.-).
DE SLL0776.
OS SYNECHOCYSTIS SP. (SPRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE: 96127529.
RX KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA RES. 2:153-166(1995).
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

-----
DR EMBL: D64005; G1006577; -.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM: PF00069; pkinase; 1.
KW HYPOTHETICAL PROTEIN; TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE;
KW ATP-BINDING.
FT DOMAIN 9 271 PROTEIN KINASE.
FT NP_BIND 15 23 ATP (BY SIMILARITY).
FT BINDING 40 40 ATP (BY SIMILARITY).
FT ACT_SITE 136 136 BY SIMILARITY.
SQ SEQUENCE 505 AA; 55213 MW; C7A77499 CRC32;

Query Match 43.1%; Score 62; DB 1; Length 505;
Best Local Similarity 47.1%; Pred. No. 1.44e+00;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db      84 FSDNDFYLVQEIYQGV 100
QY      1 YNNDKFYLSNGRIQAV 17

RESULT 13
ID PODK_MESOR STANDARD; PRT; 949 AA.
AC Q42910; Q42911;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE PYRUVATE, PHOSPHATE KINASE PRECURSOR (EC 2.7.9.1)
DE (PYRUVATE, ORTHOPHOSPHATE DIKINASE).
GN PPD OR PPDK.
OS MESEMBRYANTHEMUM CRYSTALLINUM (COMMON ICE PLANT).
OC EUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC CARYOPHYLLIDAE; CARYOPHYLLALES; ALZACEAE; MESEMBRYANTHEMUM.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LEAF;
RX MEDLINE: 95375662.
RA FISLTHALER B., MEYER G., BOHNERT H.J., SCHMITT J.M.;
RT "Age-dependent induction of pyruvate, orthophosphate dikinase in
RT Mesembryanthemum crystallinum L.";
RL PLANTA 196:492-500(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA FISLTHALER B., MEYER G., SCHMITT J.M.;
RT "The gene encoding pyruvate, orthophosphate dikinase from
RT Mesembryanthemum crystallinum L. has a long intron in the 5'
RT untranslated region.";
RL (IN) PLANT GENE REGISTRY PGR96-010.
CC -!- FUNCTION: FORMATION OF PHOSPHENOLPYRUVATE, WHICH IS THE
CC PRIMARY ACCEPTOR OF CO2 IN C4 AND SOME CRASSULACEAN ACID
CC METABOLISM PLANTS.

```





**This Page Blank (uspto)**



RX MEDLINE: 92307745.  
RA KENTER M., OTTING N., ANHOLTS J., LEUNISSEN J., JONKER M.,  
RA BONTROP R.E.;  
RT "Evolutionary relationships among the primate Mhc-DQA1 and DQA2  
alleles."  
RL IMMUNOGENETICS 36:71-78(1992).  
DR EMBL; M76190; G176843; -  
DR PFAM; PF00993; MHC\_II\_alpha; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 62  
SQ SEQUENCE 62 AA; 7247 MW; 1C603C2A CRC32;

Query Match 61.1%; Score 88; DB 7; Length 62;  
Best Local Similarity 47.4%; Pred. No. 2.21e-05;  
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30  
:: ||||: ||::||:|  
Qy 1 YNDEKFYLSNGRIQAVRC 19

RESULT 3  
ID O19630 PRELIMINARY; PRT; 40 AA.  
AC O19630;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE MHC CLASS II ANTIGEN (FRAGMENT).  
CN HLA-DOA1\*0502  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 93127145.  
RA MARSH S.G.E., BODOMER J.G.;  
RT "HLA class II nucleotide sequences, 1992."  
RL TISSUE ANTIGENS 40:229-243(1992).  
RN [2]  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 40  
SQ SEQUENCE 40 AA; 4612 MW; 94F266A5 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 40;  
Best Local Similarity 42.1%; Pred. No. 7.19e-04;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 18 FDGDEQFYVDLGRKETVWC 36  
:: ||||: ||::||:|  
Qy 1 YNDEKFYLSNGRIQAVRC 19

RESULT 4  
ID Q30996 PRELIMINARY; PRT; 62 AA.  
AC Q30996;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE MHC-DOA1\*04 (FRAGMENT).  
CN MHC DO-ALPHA 1  
OS PAN TROGLODYTES (CHIMPANZEE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; PAN.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CORRY; TISSUE-BLOOD;  
RX MEDLINE: 92307745.

RA KENTER M., OTTING N., ANHOLTS J., LEUNISSEN J., JONKER M.,  
RA BONTROP R.E.;  
RT "Evolutionary relationships among the primate Mhc-DQA1 and DQA2  
alleles."  
RL IMMUNOGENETICS 36:71-78(1992).  
DR EMBL; M76191; G176845; -  
DR PFAM; PF00993; MHC\_II\_alpha; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 62  
SQ SEQUENCE 62 AA; 7231 MW; 52CEC750 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 62;  
Best Local Similarity 42.1%; Pred. No. 7.19e-04;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 12 FDGDEQFYVDLGRKETVWC 30  
:: ||||: ||::||:|  
Qy 1 YNDEKFYLSNGRIQAVRC 19

RESULT 5  
ID O19765 PRELIMINARY; PRT; 74 AA.  
AC O19765;  
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE MHC CLASS II ANTIGEN HLA-DOA1 (FRAGMENT).  
CN HLA-DOA1.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PERIPHERAL BLOOD;  
RA LESTER S., LAHAM N., DOMSE T., MCCLUSKEY J.;  
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U97555; G2121263; -  
DR PFAM; PF00993; MHC\_II\_alpha; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 74  
SQ SEQUENCE 74 AA; 8483 MW; 201AF66D CRC32;

Query Match 56.3%; Score 81; DB 7; Length 74;  
Best Local Similarity 42.1%; Pred. No. 7.19e-04;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 18 FDGDEQFYVDLGRKETVWC 36  
:: ||||: ||::||:|  
Qy 1 YNDEKFYLSNGRIQAVRC 19

RESULT 6  
ID P79553 PRELIMINARY; PRT; 81 AA.  
AC P79553;  
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)  
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE MHC CLASS II ANTIGEN (FRAGMENT).  
CN HLA-DQ.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TREJAUT J.A., GREVILLE W.D., DUNCKLEY H.;  
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U85035; G1814392; -  
DR PFAM; PF00993; MHC\_II\_alpha; 1.  
KW MHC.  
FT NON\_TER 1  
FT NON\_TER 81

```
SQ SEQUENCE 81 AA: 9270 MW: 35141696 CRC32:
Query Match 56.3%; Score 81; DB 7; Length 81;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 24 FDGDEQFYVDLGRKETVWC 42
:: |||:: ||::|:|
QY 1 YNDEKEYLSNGRIQAVRC 19

RESULT 7
ID Q30125 PRELIMINARY; PRT; 86 AA.
AC Q30125;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-D ALPHA MRNA (HAPLOTYPE DR3), PARTIAL CDS
DE (FRAGMENT).
GN HLA-DRA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88116708.
RA HURLEY C.K., GREGGENSEN P., STEINER N., BELL J., HARTZMAN R., NEPOM G.,
RA SILVER J., JOHNSON A.H.;
RT "Polymorphism of the HLA-D region in American blacks. A DR3 haplotype
RT generated by recombination.";
RL J. IMMUNOL. 140:885-892(1988).
DR EMBL; M18935; G188261; -.
DR PFAM; PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 86 86
SQ SEQUENCE 86 AA: 9796 MW: E7124639 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 86;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGDEQFYVDLGRKETVWC 47
:: |||:: ||::|:|
QY 1 YNDEKEYLSNGRIQAVRC 19

RESULT 8
ID O19522 PRELIMINARY; PRT; 169 AA.
AC O19522;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II ANTIGEN DO-ALPHA-1 CHAIN (FRAGMENT).
GN DO-A1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA JAVAUUX F., MOLLET I., GEBUHRER L.;
RL SUBMITTED (JUL-1997) TO ENBL/GENEBANK/DBBJ DATA BANKS.
DR EMBL; AF012768; G2394011; -.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 169 169
SQ SEQUENCE 169 AA: 19107 MW: 68A6EB34 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 169;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 18 FDGDEQFYVDLGRKETVWC 36
:: |||:: ||::|:|
QY 1 YNDEKEYLSNGRIQAVRC 19

RESULT 9
ID Q30101 PRELIMINARY; PRT; 220 AA.
AC Q30101;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA (FRAGMENT).
GN HLA-DQA1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88199016.
RA LIU C.P., BACH F.H., WU S.K.;
RT "Molecular studies of a rare DR2/LD-5a/DO*3 HLA class II haplotype.
RT Multiple genetic mechanisms in the generation of polymorphic HLA
RT class II genes.";
RL J. IMMUNOL. 140:3631-3639(1988).
DR EMBL; M20506; G188207; -.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 220 220
SQ SEQUENCE 220 AA: 24370 MW: 079FD0F1 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 220;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 17 FDGDEQFYVDLGRKETVWC 35
:: |||:: ||::|:|
QY 1 YNDEKEYLSNGRIQAVRC 19

RESULT 10
ID O19705 PRELIMINARY; PRT; 226 AA.
AC O19705;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA CHAIN (FRAGMENT).
GN HLA-DQA1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97083137.
RA YASUNAGA S., KIMURA A., HAMAGUCHI K., RONNINGEN K.S., SASAZUKI T.;
RT "Different contribution of HLA-DR and -DQ genes in susceptibility and
RT resistance to insulin-dependent diabetes mellitus (IDDM).";
RL TISSUE ANTIGENS 47:37-48(1996).
DR EMBL; L34091; G619826; -.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1 1
FT NON_TER 226 226
SQ SEQUENCE 226 AA: 24984 MW: 25E83AAB CRC32;

Query Match 56.3%; Score 81; DB 7; Length 226;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGDEQFYVDLGRKETVWC 47
:: |||:: ||::|:|
```

```
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 11
ID O19706 PRELIMINARY: PRT: 226 AA.
AC O19706;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA CHAIN (FRAGMENT).
GN HLA-DQAL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97083137.
RA YASUNAGA S., KIMURA A., HAMAGUCHI K., RONNINGEN K.S., SASAZUKI T.;
RT "Different contribution of HLA-DR and -DQ genes in susceptibility and
RT resistance to insulin-dependent diabetes mellitus (IDDM).";
RL TISSUE ANTIGENS 47:37-48(1996).
DR EMBL: L34092; G619828; -.
DR PFAM: PF00047; ig: 1.
DR PFAM: PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 226
SQ SEQUENCE 226 AA; 24984 MW; 25E83AAB CRC32;

Query Match 56.3%; Score 81; DB 7; Length 226;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGEQFYVDLGRKETVWC 47
:: || ||: ||: ||: ||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 12
ID Q30072 PRELIMINARY: PRT: 226 AA.
AC Q30072;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA CHAIN (FRAGMENT).
GN HLA-DQAL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97083137.
RA YASUNAGA S., KIMURA A., HAMAGUCHI K., RONNINGEN K.S., SASAZUKI T.;
RT "Different contribution of HLA-DR and -DQ genes in susceptibility and
RT resistance to insulin-dependent diabetes mellitus (IDDM).";
RL TISSUE ANTIGENS 47:37-48(1996).
DR EMBL: L34093; G619830; -.
DR PFAM: PF00047; ig: 1.
DR PFAM: PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 226
SQ SEQUENCE 226 AA; 25000 MW; 0AFF978A CRC32;

Query Match 56.3%; Score 81; DB 7; Length 226;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGEQFYVDLGRKETVWC 47
:: || ||: ||: ||: ||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 13
ID Q30073 PRELIMINARY: PRT: 226 AA.
AC Q30073;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA CHAIN (FRAGMENT).
GN HLA-DQAL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97083137.
RA YASUNAGA S., KIMURA A., HAMAGUCHI K., RONNINGEN K.S., SASAZUKI T.;
RT "Different contribution of HLA-DR and -DQ genes in susceptibility and
RT resistance to insulin-dependent diabetes mellitus (IDDM).";
RL TISSUE ANTIGENS 47:37-48(1996).
DR EMBL: L34094; G619832; -.
DR PFAM: PF00047; ig: 1.
DR PFAM: PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 226
SQ SEQUENCE 226 AA; 25017 MW; DC29835A CRC32;

Query Match 56.3%; Score 81; DB 7; Length 226;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGEQFYVDLGRKETVWC 47
:: || ||: ||: ||: ||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 14
ID Q30071 PRELIMINARY: PRT: 226 AA.
AC Q30071;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II HLA-DQ-ALPHA CHAIN (FRAGMENT).
GN HLA-DQAL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97083137.
RA YASUNAGA S., KIMURA A., HAMAGUCHI K., RONNINGEN K.S., SASAZUKI T.;
RT "Different contribution of HLA-DR and -DQ genes in susceptibility and
RT resistance to insulin-dependent diabetes mellitus (IDDM).";
RL TISSUE ANTIGENS 47:37-48(1996).
DR EMBL: L34090; G619824; -.
DR PFAM: PF00047; ig: 1.
DR PFAM: PF00993; MHC_II_alpha; 1.
KW MHC.
FT NON_TER 1
FT NON_TER 226
SQ SEQUENCE 226 AA; 25033 MW; A6F0425C CRC32;

Query Match 56.3%; Score 81; DB 7; Length 226;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 29 FDGEQFYVDLGRKETVWC 47
:: || ||: ||: ||: ||
QY 1 YNDEKFYLSNGRIQAVRC 19

RESULT 15
ID Q30085 PRELIMINARY: PRT: 254 AA.
AC Q30085;
```

Mon Aug 30 10:55:02 1999

```

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE MHC CLASS II DQ ALPHA PRECURSOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89379293.
RA JONSSON A.K., ANDERSSON L., RASK L.;
RT "Complete sequences of DQa1 and DQb1 cDNA clones corresponding to the
RL DQa1 specificity."
RL IMMUNOGENETICS 30:232-234(1989).
DR EMBL; M26041; G307251; -.
DR PFAM; PF00047; Ig; 1.
DR PFAM; PF00993; MHC_II_alpha; 1.
KW SIGNAL; MHC.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 254 POTENTIAL.
SQ SEQUENCE 254 AA; 27836 MW; 6C71B522 CRC32;

Query Match 56.3%; Score 81; DB 7; Length 254;
Best Local Similarity 42.1%; Pred. No. 7.19e-04;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 52 FDGDEQFYVDLGRKETVWC 70
QY 1 YNDEKFYLSNGRIQAVRC 19

```

Search completed: Mon Aug 23 13:03:11 1999  
Job time : 20 secs.

**This Page Blank (uspto)**



\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:35:05 1999; MasPar time 35.97 Seconds  
540.412 Million cell updates/sec  
Tabular output not generated.

```
>US-09-049-696-41
Title:
Description:
Perfect Score: 914
Sequence: 1 MGPFKSSVFILILHLEGAL.....GIHLKIMKWKIGELQSLA 914
```

Scoring table: TABLE unitprotable  
Gap 60

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database:

a-geneseq35

1:part1	2:part2	3:part3	4:part4	5:part5	6:part6	7:part7
8:part8	9:part9	10:part10	11:part11	12:part12	13:part13	
14:part14	15:part15	16:part16	17:part17	18:part18		
19:part19	20:part20	21:part21	22:part22	23:part23		
24:part24	25:part25	26:part26	27:part27	28:part28		
29:part29	30:part30	31:part31	32:part32	33:part33		
34:part34	35:part35	36:part36	37:part37	38:part38		
39:part39						

Statistics: Mean 3.243; Variance 0.732; scale 4.428

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	228	24.9	228 20	W06548	Human colon specific	0.00e+00
2	228	24.9	228 29	W46879	Protein sequence enco	0.00e+00
3	228	7	15 20	W04842	Self epitope of desmo	5.26e+01
4	7	0.8	19 16	R33961	Peptide combining wit	5.26e+01
5	7	0.8	21 10	R47000	LAM Blast-1 with N-ac	5.26e+01
6	7	0.8	21 16	R93962	Peptide combining wit	5.26e+01
7	7	0.8	26 2	R07135	Peptide with hairpin	5.26e+01
8	7	0.8	30 16	R93960	Peptide combining wit	5.26e+01
9	7	0.8	37 28	W43445	Scorpion toxin-relate	5.26e+01
10	7	0.8	119 29	W42471	NEWM humanised heavy	5.26e+01
11	7	0.8	120 17	R98109	Murine anti-Protein C	5.26e+01
12	7	0.8	133 26	W58794	Maize r12 gene partia	5.26e+01
13	7	0.8	139 17	R98107	Murine anti-Protein C	5.26e+01
14	7	0.8	182 35	W76251	Human TACE-like prote	5.26e+01
15	7	0.8	194 27	W35857	Human CD48 for use in	5.26e+01
16	7	0.8	211 17	R98658	N.meningitidis M978f	5.26e+01

90	6	0.7	94	39	W90124	Human chemokine MIG-b	6.28e+02	163	6	0.7	110	6	R33314	Variant IgE - mutant	6.28e+02
91	6	0.7	94	30	W57044	Human chemokine alpha	6.28e+02	164	6	0.7	110	6	R32040	Variant IgE - mutant	6.28e+02
92	6	0.7	94	25	W28512	Product of clone H174	6.28e+02	165	6	0.7	110	6	R33331	Variant IgE - mutant	6.28e+02
93	6	0.7	94	25	W28513	Product of clone H174	6.28e+02	166	6	0.7	110	6	R33682	Variant IgE - mutant	6.28e+02
94	6	0.7	95	34	W69989	Primate CXC chemokine	6.28e+02	167	6	0.7	111	6	R33313	Variant IgE - mutant	6.28e+02
95	6	0.7	97	34	W71021	Mus musculus Ggamma3	6.28e+02	168	6	0.7	112	1	P90006	Immunoglobulin E comp	6.28e+02
96	6	0.7	98	17	R67892	Mengovirus leader-ICM	6.28e+02	169	6	0.7	113	13	R71833	Almonaria alternata	6.28e+02
97	6	0.7	98	31	W27948	Amino acid sequence o	6.28e+02	170	6	0.7	116	22	W20660	H. pylori cytoplasmic	6.28e+02
98	6	0.7	103	29	W52509	H. pylori ORF 02ael16	6.28e+02	171	6	0.7	117	31	W58504	Human heavy chain sub	6.28e+02
99	6	0.7	105	29	W52233	Antibody L02-14-VL ch	6.28e+02	172	6	0.7	117	1	P32036	Sequence encoded in t	6.28e+02
100	6	0.7	109	30	W52839	Secreted protein enco	6.28e+02	173	6	0.7	117	1	P30153	Sequence of hepatitis	6.28e+02
101	6	0.7	110	6	R33327	Variant IgE - mutant	6.28e+02	174	6	0.7	119	29	W42465	Interleukin-5 humanis	6.28e+02
102	6	0.7	110	6	R33693	Variant IgE - mutant	6.28e+02	175	6	0.7	119	18	R98488	Humanised 2B6 antibod	6.28e+02
103	6	0.7	110	6	R33753	Variant IgE - mutant	6.28e+02	176	6	0.7	119	12	R66299	Human immunoglobulin	6.28e+02
104	6	0.7	110	12	R65464	T-cell receptor V-bet	6.28e+02	177	6	0.7	119	29	W26783	Interleukin-5 humanis	6.28e+02
105	6	0.7	110	6	R32024	Variant IgE - mutant	6.28e+02	178	6	0.7	119	29	W26784	Interleukin-5 humanis	6.28e+02
106	6	0.7	110	6	R33745	Variant IgE - mutant	6.28e+02	179	6	0.7	120	16	R32088	CDR-grafted anti-RSV	6.28e+02
107	6	0.7	110	6	R33678	Variant IgE - mutant	6.28e+02	180	6	0.7	120	16	R27086	Human Cos VH region.	6.28e+02
108	6	0.7	110	6	R33683	Variant IgE - mutant	6.28e+02	181	6	0.7	121	25	W27552	Human Ab heavy chain	6.28e+02
109	6	0.7	110	6	R33743	Variant IgE - mutant	6.28e+02	182	6	0.7	123	5	R26055	Human galanin.	6.28e+02
110	6	0.7	110	6	R33691	Variant IgE - mutant	6.28e+02	183	6	0.7	123	5	R27115	Sequence of human pre	6.28e+02
111	6	0.7	110	6	R33318	Variant IgE - mutant	6.28e+02	184	6	0.7	125	8	R45157	Botrocetin beta subun	6.28e+02
112	6	0.7	110	6	R33744	Variant IgE - mutant	6.28e+02	185	6	0.7	126	21	W11635	Human anti-RSV monoc	6.28e+02
113	6	0.7	110	6	R32030	Variant IgE - mutant	6.28e+02	186	6	0.7	128	21	W04211	Streptomyces avidinil	6.28e+02
114	6	0.7	110	6	R32041	Variant IgE - mutant	6.28e+02	187	6	0.7	128	21	W20127	H. pylori cytoplasmic	6.28e+02
115	6	0.7	110	6	R33747	Variant IgE - mutant	6.28e+02	188	6	0.7	128	7	R34722	Core streptavidin.	6.28e+02
116	6	0.7	110	6	R33748	Variant IgE - mutant	6.28e+02	189	6	0.7	128	27	W29308	Recombinant Core-stre	6.28e+02
117	6	0.7	110	6	R32035	Variant IgE - mutant	6.28e+02	190	6	0.7	131	32	W26065	Human stem cell anti	6.28e+02
118	6	0.7	110	12	R65465	T-cell receptor V-bet	6.28e+02	191	6	0.7	131	33	W69214	Osteoclast inhibitor	6.28e+02
119	6	0.7	110	6	R32025	Variant IgE - mutant	6.28e+02	192	6	0.7	131	17	R33017	Haematopoietic stem c	6.28e+02
120	6	0.7	110	6	R33684	Variant IgE - mutant	6.28e+02	193	6	0.7	131	25	W24847	Human stem cell anti	6.28e+02
121	6	0.7	110	6	R32027	Variant IgE - mutant	6.28e+02	194	6	0.7	132	1	P90157	Sequence of hepatitis	6.28e+02
122	6	0.7	110	6	R33692	Variant IgE - mutant	6.28e+02	195	6	0.7	132	1	P92040	Sequence encoded in t	6.28e+02
123	6	0.7	110	6	R33315	Variant IgE - mutant	6.28e+02	196	6	0.7	132	37	W75070	Human secreted protei	6.28e+02
124	6	0.7	110	6	R33317	Variant IgE - mutant	6.28e+02	197	6	0.7	132	3	R15478	Modified, truncated C	6.28e+02
125	6	0.7	110	6	R33316	Variant IgE - mutant	6.28e+02	198	6	0.7	139	13	R67658	Anti-human IL-6 china	6.28e+02
126	6	0.7	110	6	R33320	Variant IgE - mutant	6.28e+02	199	6	0.7	139	13	R76759	Anti-human IL-6 china	6.28e+02
127	6	0.7	110	6	R33323	Variant IgE - mutant	6.28e+02	200	6	0.7	139	7	R38315	Sequence of the VH of	6.28e+02
128	6	0.7	110	6	R32034	Variant IgE - mutant	6.28e+02	201	6	0.7	141	13	R70192	Humanized antibody 3B	6.28e+02
129	6	0.7	110	6	R33329	Variant IgE - mutant	6.28e+02	202	6	0.7	142	2	P81064	Sequence of rhinoviru	6.28e+02
130	6	0.7	110	12	R65466	T-cell receptor V-bet	6.28e+02	203	6	0.7	148	27	W62339	Monoclonal antibody 1	6.28e+02
131	6	0.7	110	6	R33689	Variant IgE - mutant	6.28e+02	204	6	0.7	148	20	R99844	Human natural killer	6.28e+02
132	6	0.7	110	6	R33688	Variant IgE - mutant	6.28e+02	205	6	0.7	153	1	P90951	Herpes Simplex virus-	6.28e+02
133	6	0.7	110	6	R32028	Variant IgE - mutant	6.28e+02	206	6	0.7	155	35	W74475	Amino acid sequence o	6.28e+02
134	6	0.7	110	6	R32028	Variant IgE - mutant	6.28e+02	207	6	0.7	157	32	W58594	Mouse protein L256-11	6.28e+02
135	6	0.7	110	6	R33750	Variant IgE - mutant	6.28e+02	208	6	0.7	159	27	W29311	Streptavidin protein	6.28e+02
136	6	0.7	110	6	R32039	Variant IgE - mutant	6.28e+02	209	6	0.7	159	27	W29312	Streptavidin protein	6.28e+02
137	6	0.7	110	6	R33679	Variant IgE - mutant	6.28e+02	210	6	0.7	159	27	W29319	Streptavidin protein	6.28e+02
138	6	0.7	110	6	R33751	Variant IgE - mutant	6.28e+02	211	6	0.7	159	27	W29313	Streptavidin protein	6.28e+02
139	6	0.7	110	6	R33321	Variant IgE - mutant	6.28e+02	212	6	0.7	159	27	W29320	Streptavidin protein	6.28e+02
140	6	0.7	110	6	R33319	Variant IgE - mutant	6.28e+02	213	6	0.7	159	27	W29316	Streptavidin protein	6.28e+02
141	6	0.7	110	6	R32023	Variant IgE - mutant	6.28e+02	214	6	0.7	159	27	W29315	Streptavidin protein	6.28e+02
142	6	0.7	110	6	R32022	Variant IgE - mutant	6.28e+02	215	6	0.7	159	27	W29309	Streptavidin protein	6.28e+02
143	6	0.7	110	6	R33752	Variant IgE - mutant	6.28e+02	216	6	0.7	159	27	W29318	Streptavidin protein	6.28e+02
144	6	0.7	110	6	R33681	Variant IgE - mutant	6.28e+02	217	6	0.7	159	27	W29314	Streptavidin protein	6.28e+02
145	6	0.7	110	6	R33680	Variant IgE - mutant	6.28e+02	218	6	0.7	159	27	W29310	Streptavidin protein	6.28e+02
146	6	0.7	110	6	R33690	Variant IgE - mutant	6.28e+02	219	6	0.7	159	27	W29317	Streptavidin protein	6.28e+02
147	6	0.7	110	6	R33686	Variant IgE - mutant	6.28e+02	220	6	0.7	160	1	P93531	Mature streptavidin a	6.28e+02
148	6	0.7	110	6	R33687	Variant IgE - mutant	6.28e+02	221	6	0.7	163	1	P80160	Biosynthetic protein	6.28e+02
149	6	0.7	110	6	R32031	Variant IgE - mutant	6.28e+02	222	6	0.7	165	1	R05422	Human ventricular myo	6.28e+02
150	6	0.7	110	6	R32038	Variant IgE - mutant	6.28e+02	223	6	0.7	166	27	W35329	Human cdc35B vaccinia	6.28e+02
151	6	0.7	110	6	R32026	Variant IgE - mutant	6.28e+02	224	6	0.7	166	21	W13146	Human soluble CD23 is	6.28e+02
152	6	0.7	110	6	R33328	Variant IgE - mutant	6.28e+02	225	6	0.7	167	21	W13148	Human soluble CD23 is	6.28e+02
153	6	0.7	110	6	R33746	Variant IgE - mutant	6.28e+02	226	6	0.7	167	34	W38557	Streptococcus pneumo	6.28e+02
154	6	0.7	110	6	R33749	Variant IgE - mutant	6.28e+02	227	6	0.7	173	4	R22666	Collagen type protein	6.28e+02
155	6	0.7	110	6	R32037	Variant IgE - mutant	6.28e+02	228	6	0.7	181	30	W44833	Rat p26 protein sequ	6.28e+02
156	6	0.7	110	6	R32036	Variant IgE - mutant	6.28e+02	229	6	0.7	181	30	W44835	Mouse p26 protein.	6.28e+02
157	6	0.7	110	6	R33330	Variant IgE - mutant	6.28e+02	230	6	0.7	182	2	P70492	Streptavidin sequence	6.28e+02
158	6	0.7	110	6	R33304	IgE Fc epsilon 3.	6.28e+02	231	6	0.7	183	31	W59217	S. avidinii streptavi	6.28e+02
159	6	0.7	110	6	R33685	Variant IgE - mutant	6.28e+02	232	6	0.7	183	29	W41225	Priort protein binding	6.28e+02
160	6	0.7	110	6	R32032	Variant IgE - mutant	6.28e+02	233	6	0.7	183	31	W59216	S. avidinii streptavi	6.28e+02
161	6	0.7	110	6	R32033	Variant IgE - mutant	6.28e+02	234	6	0.7	183	31	W59218	S. avidinii streptavi	6.28e+02
162	6	0.7	110	6	R33324	Variant IgE - mutant	6.28e+02	235	6	0.7	183	1	P93530	Streptavidin protein	6.28e+02

236	6	0.7	183	27	W29306	Wild-type streptavidin	6.28e+02	309	263	12	R62554	Activated modified hu	6.28e+02
237	6	0.7	183	3	P60625	Sequence of a strepta	6.28e+02	310	263	1	P90507	Sequence of an epitop	6.28e+02
238	6	0.7	183	9	R44491	Streptavidin gene.	6.28e+02	311	263	1	P93565	Sequence of an epitop	6.28e+02
239	6	0.7	183	33	W60978	Streptococcus pneumon	6.28e+02	312	264	5	R29638	pCTD ORF 3.	6.28e+02
240	6	0.7	189	36	W79541	Adenovirus serotype 5	6.28e+02	313	265	38	W89987	Expressed antigen for	6.28e+02
241	6	0.7	190	15	R85584	Fc(epsilon) CH2'-CH4	6.28e+02	314	265	28	W35318	Human cdc25B302-566 p	6.28e+02
242	6	0.7	188	17	R88660	N.meningitidis S3032	6.28e+02	315	265	27	W35324	Human cdc25B302-566 M	6.28e+02
243	6	0.7	198	17	R88652	N.meningitidis IM2169	6.28e+02	316	265	29	W55669	H. pylori ORF hp5p152	6.28e+02
244	6	0.7	199	17	R88661	N.meningitidis B283 T	6.28e+02	317	265	27	W35323	Human cdc25B302-566 M	6.28e+02
245	6	0.7	201	36	W79540	Adenovirus serotype 2	6.28e+02	318	265	27	W35322	Human cdc25B302-566 M	6.28e+02
246	6	0.7	201	37	W35325	Human cdc25B356-566 p	6.28e+02	319	265	27	W35319	Human Gly-ile-cdc25B3	6.28e+02
247	6	0.7	201	15	R85589	Fc(epsilon) CH2'-CH4	6.28e+02	320	267	38	W89860	Antigen 1 from cluste	6.28e+02
248	6	0.7	204	24	W14378	Streptococcus pneumon	6.28e+02	321	269	38	W89860	Genomic sequence of h	6.28e+02
249	6	0.7	207	31	W28088	Staphylococcus aureus	6.28e+02	322	271	27	W35320	Human Ile-Glu-Gly-Arg	6.28e+02
250	6	0.7	209	34	W64466	Human secreted protei	6.28e+02	323	271	1	R04217	Human 32K ASP encoded	6.28e+02
251	6	0.7	209	1	P90010	Residues 340-547 of t	6.28e+02	324	271	1	R04212	Human 32K alveolar su	6.28e+02
252	6	0.7	211	12	R62661	Hepatitis Type-C viru	6.28e+02	325	274	1	R06836	Intercellular adhesio	6.28e+02
253	6	0.7	211	27	W35321	Human cdc25B356-566 p	6.28e+02	326	278	33	W57572	Homo sapiens B268 seq	6.28e+02
254	6	0.7	211	27	W35327	Human cdc25B356-566 M	6.28e+02	327	280	37	W74824	Human secreted protei	6.28e+02
255	6	0.7	213	16	R91042	V8 mature protease (a	6.28e+02	328	283	35	W64777	C. magnoliae carbonyl	6.28e+02
256	6	0.7	214	16	R91043	V8 mature protease (a	6.28e+02	329	291	32	W48858	Human retinaldehyde b	6.28e+02
257	6	0.7	215	8	R45007	Sequence encoded by a	6.28e+02	330	292	6	R29312	Cyclin D3 protein.	6.28e+02
258	6	0.7	215	16	R91044	V8 mature protease (a	6.28e+02	331	292	8	R44804	Human cyclin D3	6.28e+02
259	6	0.7	218	15	R85583	Fc(epsilon) CH2'-CH4	6.28e+02	332	295	38	W82588	Human Arg-1622 protei	6.28e+02
260	6	0.7	220	17	R67890	Del-gp120-L fusion pr	6.28e+02	333	295	34	W49082	Homo sapiens Sdr-5 pr	6.28e+02
261	6	0.7	222	10	R53058	Human mature cathepsi	6.28e+02	334	295	32	W48861	Human retinaldehyde b	6.28e+02
262	6	0.7	223	37	W81767	Bovine trypsin	6.28e+02	335	295	32	W37814	Murine secreted apopt	6.28e+02
263	6	0.7	224	10	R53637	Bovine trypsin.	6.28e+02	336	295	32	W37943	Amino acid sequence o	6.28e+02
264	6	0.7	227	22	W24561	H. pylori cytoplasmic	6.28e+02	337	295	37	W75102	Human secreted protei	6.28e+02
265	6	0.7	227	22	W20435	H. pylori cytoplasmic	6.28e+02	338	296	37	W75160	Human secreted protei	6.28e+02
266	6	0.7	228	22	W20224	H. pylori transporter	6.28e+02	339	300	1	P83022	Human IgE binding fac	6.28e+02
267	6	0.7	230	10	R53638	Bovine trypsinogen.	6.28e+02	340	302	22	W20681	H. pylori cytoplasmic	6.28e+02
268	6	0.7	232	16	R89430	Trypsin-like enzyme.	6.28e+02	341	304	5	R25291	Acetyl xylan esterase	6.28e+02
269	6	0.7	233	23	W13585	Batten disease CLN3 m	6.28e+02	342	307	22	W24626	H. pylori cytoplasmic	6.28e+02
270	6	0.7	234	23	W13584	Batten disease CLN3	6.28e+02	343	307	22	W20285	H. pylori cytoplasmic	6.28e+02
271	6	0.7	234	8	R45009	Sequence encoded by a	6.28e+02	344	315	19	W03679	G-protein coupled hum	6.28e+02
272	6	0.7	235	22	W18648	Human ubiquitin gene	6.28e+02	345	315	15	R85582	Fc(epsilon) CH2'-CH4	6.28e+02
273	6	0.7	237	22	W10661	Recombinant endoglyco	6.28e+02	346	315	16	R48707	G-protein coupled hum	6.28e+02
274	6	0.7	239	22	W20567	H. pylori cytoplasmic	6.28e+02	347	315	16	R28809	Rhipecephalus appendi	6.28e+02
275	6	0.7	241	22	W20668	H. pylori transporter	6.28e+02	348	316	22	W20864	H. pylori cytoplasmic	6.28e+02
276	6	0.7	244	31	W51009	Pig lung carbonyl red	6.28e+02	349	317	22	W14320	T. neopolitana xyna d	6.28e+02
277	6	0.7	246	6	R29894	HCV NS4-NS5 peptide O	6.28e+02	350	320	21	W13143	Sequence containing h	6.28e+02
278	6	0.7	246	6	R29893	HCV NS4-NS5 peptide O	6.28e+02	351	321	21	W13142	Sequence containing h	6.28e+02
279	6	0.7	246	35	W64260	Human amyloid beta-pr	6.28e+02	352	321	2	P90105	IgE binding factor.	6.28e+02
280	6	0.7	247	20	W08475	Porcine trypsinogen.	6.28e+02	353	321	1	P90120	Human lymphocyte rece	6.28e+02
281	6	0.7	248	3	P60442	Plasmid pASpcg-SV(10)	6.28e+02	354	321	1	P90367	Peptide sequence of m	6.28e+02
282	6	0.7	248	3	P70663	35kd pulmonary surfac	6.28e+02	355	321	1	P81230	Fc gamma receptor	6.28e+02
283	6	0.7	248	3	P60665	Sequence of human alv	6.28e+02	356	321	8	R42053	IgE binding factor.	6.28e+02
284	6	0.7	248	3	P60666	Genomic sequence of h	6.28e+02	357	321	3	P81172	Sequence encoded by a	6.28e+02
285	6	0.7	248	1	R04216	Human 32K ASP encoded	6.28e+02	358	321	1	P82073	Recombinant Fc-epsilo	6.28e+02
286	6	0.7	248	1	P80594	Sequence deduced from	6.28e+02	359	321	1	P81163	Low affinity Fc-epsilo	6.28e+02
287	6	0.7	248	2	R05092	Gene product of vecto	6.28e+02	360	321	1	P81112	Human IgE binding fac	6.28e+02
288	6	0.7	248	3	P60441	Plasmid pASpc-SV(10)	6.28e+02	361	322	4	R23599	Recombinant hematopei	6.28e+02
289	6	0.7	248	1	R06331	Human alveolar surfac	6.28e+02	362	324	15	R83559	Fc(epsilon) CH2'-CH4	6.28e+02
290	6	0.7	248	1	R04215	Human 32K ASP encoded	6.28e+02	363	325	16	R83582	CH2 to CH4 of human I	6.28e+02
291	6	0.7	249	38	W89822	Protein encoded by cl	6.28e+02	364	325	13	R77241	Human IgE FC chain (a	6.28e+02
292	6	0.7	250	21	W14446	Carb gene product.	6.28e+02	365	325	13	R75225	Human IgE FC chain (a	6.28e+02
293	6	0.7	251	28	W37358	HisF protein involved	6.28e+02	366	331	19	W02626	Human native c-Jun pr	6.28e+02
294	6	0.7	251	38	R13048	Protein C heavy chain	6.28e+02	367	331	37	W73190	c-Jun protein sequenc	6.28e+02
295	6	0.7	260	38	W86631	Secreted protein enco	6.28e+02	368	331	10	R53645	c-jun gene product.	6.28e+02
296	6	0.7	260	39	W87703	A human serine protea	6.28e+02	369	331	19	W02627	Human modified c-Jun	6.28e+02
297	6	0.7	261	6	R30723	Protein C heavy chain	6.28e+02	370	331	13	R70810	c-jun oncoprotein.	6.28e+02
298	6	0.7	262	2	R12194	Human protein C catal	6.28e+02	371	331	35	W70337	c-jun oncoprotein.	6.28e+02
299	6	0.7	262	2	R12196	Human protein C catal	6.28e+02	372	334	10	R53644	c-jun gene gene produ	6.28e+02
300	6	0.7	262	1	R05276	Alkaline protease enc	6.28e+02	373	334	23	W22732	Human ATP receptor.	6.28e+02
301	6	0.7	262	3	R13047	Protein C heavy chain	6.28e+02	374	334	23	W19854	Human purinergic rece	6.28e+02
302	6	0.7	262	3	R13724	Protein C heavy chain	6.28e+02	375	336	1	P82839	Human low affinity Fc	6.28e+02
303	6	0.7	262	2	R12193	Human protein C catal	6.28e+02	376	336	5	R29644	Protease from S. Aure	6.28e+02
304	6	0.7	262	2	R12195	Human protein C catal	6.28e+02	377	337	14	R76744	Human EPR-1.	6.28e+02
305	6	0.7	262	2	R11838	Human protein C catal	6.28e+02	378	341	35	W75916	Human p58 receptor pr	6.28e+02
306	6	0.7	262	3	R13723	Protein C heavy chain	6.28e+02	379	344	23	W22218	Protein encoded by pV	6.28e+02
307	6	0.7	262	3	R13726	B10 single chain T-ce	6.28e+02	380	344	16	R91033	Beta-galactosidase-V8	6.28e+02
308	6	0.7	262	2	R12192	Human protein C catal	6.28e+02	381	345	31	W41166	Metal-regulated trans	6.28e+02

382	6	0.7	345	3	R12595	Antigenic portion of	6.28e+02	6	0.7	411	2	P70278	Streptococcus pyrogen	6.28e+02
383	6	0.7	348	8	R40973	reCA gene product.	6.28e+02	6	0.7	411	16	R83399	Fusarium oxysporum en	6.28e+02
384	6	0.7	348	25	W24801	Spinocerebellar ataxi	6.28e+02	6	0.7	415	11	R56484	ScFv PRAS109 and pRAS	6.28e+02
385	6	0.7	349	4	R23079	Epo-IL-3 Flex, recomb	6.28e+02	6	0.7	416	38	W88745	Secreted protein enco	6.28e+02
386	6	0.7	353	3	R20178	P.gluAAE PGI lipase s	6.28e+02	6	0.7	416	20	W07325	Oilpalm ACP thioester	6.28e+02
387	6	0.7	354	2	R10663	Bani restriction endo	6.28e+02	6	0.7	416	30	W47034	Human cartilage gp39-	6.28e+02
388	6	0.7	355	1	R82113	La autoantigen.	6.28e+02	6	0.7	417	35	W70386	Human cartilage gp39-	6.28e+02
389	6	0.7	355	12	R64226	Human 38kDa FK-506 bi	6.28e+02	6	0.7	418	16	R89435	Amino acid sequence o	6.28e+02
390	6	0.7	357	8	R44216	Streptomyces fradiae	6.28e+02	6	0.7	419	17	R35760	Trypsin-like enzyme.	6.28e+02
391	6	0.7	357	5	R26842	Protease from S. Aure	6.28e+02	6	0.7	419	36	W72753	Protein C (PC).	6.28e+02
392	6	0.7	361	1	P90506	Sequence of an epitop	6.28e+02	6	0.7	419	18	R96214	Primary structure of	6.28e+02
393	6	0.7	361	18	R91444	Human haematopoietic	6.28e+02	6	0.7	419	18	R96214	Recombinant human D4	6.28e+02
394	6	0.7	361	4	R20816	Haematopoietic CD44 A	6.28e+02	6	0.7	421	24	W25945	Sequence encoded by a	6.28e+02
395	6	0.7	362	1	P93567	Meroprotease antigen enc	6.28e+02	6	0.7	421	31	W53121	Ob binding protein.	6.28e+02
396	6	0.7	362	1	P90419	Plasmodium falciparum	6.28e+02	6	0.7	421	31	W53122	Amino acid sequence o	6.28e+02
397	6	0.7	362	1	P93566	41 kD merozoite antig	6.28e+02	6	0.7	423	19	R93306	Bluegill sunfish sac	6.28e+02
398	6	0.7	363	38	W88563	Secreted protein enco	6.28e+02	6	0.7	423	21	W12720	PhIE gene product.	6.28e+02
399	6	0.7	365	9	R30040	xMEF2.	6.28e+02	6	0.7	423	27	W27635	Streptococcus pneumon	6.28e+02
400	6	0.7	367	38	W89051	Polypeptide fragment	6.28e+02	6	0.7	423	35	W70387	Amino acid sequence o	6.28e+02
401	6	0.7	367	1	P80291	Interleukin-2/IgE Fc	6.28e+02	6	0.7	425	8	R43132	VPI/P2A protein.	6.28e+02
402	6	0.7	367	36	W74960	Human secreted protei	6.28e+02	6	0.7	426	6	R31036	C. acidovorans stereo	6.28e+02
403	6	0.7	368	33	W51100	Streptococcus pneumon	6.28e+02	6	0.7	427	3	R15239	Fusarium oxysporum C-	6.28e+02
404	6	0.7	370	38	W88777	Polypeptide fragment	6.28e+02	6	0.7	428	8	R42950	Human IgE heavy chain	6.28e+02
405	6	0.7	370	5	R26183	MS2-pgp3D protein.	6.28e+02	6	0.7	432	22	W20733	H. pylori cell envelo	6.28e+02
406	6	0.7	372	28	W43284	N-methylamino acid ox	6.28e+02	6	0.7	433	21	W14001	Enolase protein.	6.28e+02
407	6	0.7	374	19	R93253	Cytoplasmic antiprote	6.28e+02	6	0.7	435	31	W53917	SLG-13 protein.	6.28e+02
408	6	0.7	374	2	P80277	Sequence encoded by 3	6.28e+02	6	0.7	435	3	R12972	S-Locus specific glyco	6.28e+02
409	6	0.7	378	29	W41508	Aplysia cAMP-response	6.28e+02	6	0.7	435	11	R56483	ScFv PRAS108 and pRAS	6.28e+02
410	6	0.7	379	1	P90508	Sequence of an epitop	6.28e+02	6	0.7	437	7	R36801	Rat serotonin St-B17	6.28e+02
411	6	0.7	380	1	P90512	Sequence of new polyp	6.28e+02	6	0.7	447	17	P95275	Nisin nlsk gene produ	6.28e+02
412	6	0.7	380	1	P90509	Sequence of an epitop	6.28e+02	6	0.7	449	1	R31223	Coenzyme-independent	6.28e+02
413	6	0.7	380	1	P90510	Sequence of an epitop	6.28e+02	6	0.7	450	13	R70030	UTH1 gene product.	6.28e+02
414	6	0.7	380	1	P90511	Sequence of an epitop	6.28e+02	6	0.7	453	15	R86293	Proteinoginobacter rubru	6.28e+02
415	6	0.7	381	1	P90512	Pseudomonas aureofaci	6.28e+02	6	0.7	453	17	W55081	Streptococcus pneumon	6.28e+02
416	6	0.7	387	18	R87534	Homology vector 443-8	6.28e+02	6	0.7	455	17	R89136	Bombyx mori mature LP	6.28e+02
417	6	0.7	387	21	W10726	Human dopamine D4 rec	6.28e+02	6	0.7	456	1	R04400	Tyrosine phenol-lyase	6.28e+02
418	6	0.7	387	12	R58857	Fowlpox virus (FPV) g	6.28e+02	6	0.7	456	17	R90994	Tyrosine phenol-lyase.	6.28e+02
419	6	0.7	387	26	R31305	Pseudomonas aureofaci	6.28e+02	6	0.7	457	27	W36485	Mouse TUB Form I.	6.28e+02
420	6	0.7	387	14	R75957	Human dopamine D4 rec	6.28e+02	6	0.7	460	3	R13537	Human Protein C zymog	6.28e+02
421	6	0.7	387	30	W40503	Human dopamine D4 rec	6.28e+02	6	0.7	460	1	P81104	Sequence of human pro	6.28e+02
422	6	0.7	387	35	W63402	Phenazine gene cluste	6.28e+02	6	0.7	460	3	R13623	Human Protein C zymog	6.28e+02
423	6	0.7	387	13	R70734	Human D4 dopamine rec	6.28e+02	6	0.7	460	24	W25086	Human protein C.	6.28e+02
424	6	0.7	387	5	R25335	D4 dopamine receptor.	6.28e+02	6	0.7	460	3	R13538	Human Protein C zymog	6.28e+02
425	6	0.7	387	18	R96213	Recombinant human D4	6.28e+02	6	0.7	461	3	R13584	Human protein C zymog	6.28e+02
426	6	0.7	387	9	R48948	Sequence encoded by a	6.28e+02	6	0.7	461	3	R13582	Human protein C zymog	6.28e+02
427	6	0.7	388	19	W04405	Mouse CRTAM.	6.28e+02	6	0.7	461	1	P81205	Human protein C.	6.28e+02
428	6	0.7	391	12	R71620	Murine TRK NC2 isofo	6.28e+02	6	0.7	461	2	P70855	Human protein C.	6.28e+02
429	6	0.7	392	16	R91034	Beta-galactosidase-V8	6.28e+02	6	0.7	461	3	R13997	Human protein C.	6.28e+02
430	6	0.7	397	18	R95913	Neural thread protein	6.28e+02	6	0.7	461	3	P60001	Sequence of polypepti	6.28e+02
431	6	0.7	397	15	R83041	Capsular polysacchari	6.28e+02	6	0.7	461	1	P93714	Hybrid protein of pro	6.28e+02
432	6	0.7	402	11	R56485	ScFv PRAS110 and pRAS	6.28e+02	6	0.7	461	7	R34295	Protein C.	6.28e+02
433	6	0.7	403	6	R33251	Expression vector pAP	6.28e+02	6	0.7	461	3	R13539	Human Protein C zymog	6.28e+02
434	6	0.7	403	2	R07576	Prepro-alkaline prote	6.28e+02	6	0.7	461	1	P90401	Zymogen form of human	6.28e+02
435	6	0.7	403	7	R30672	Aspergillus Cys-169/C	6.28e+02	6	0.7	461	1	P90070	Human protein C	6.28e+02
436	6	0.7	403	1	R03610	Preproalkaline protea	6.28e+02	6	0.7	461	3	R13585	Human protein C zymog	6.28e+02
437	6	0.7	403	7	R12457	Yellow koji mould pre	6.28e+02	6	0.7	461	3	R13540	Human protein C zymog	6.28e+02
438	6	0.7	403	7	R35035	Aspergillus alkali pr	6.28e+02	6	0.7	461	19	W02600	Human protein C.	6.28e+02
439	6	0.7	403	1	R06275	Prepro-alkaline prote	6.28e+02	6	0.7	461	3	R13081	Human protein C.	6.28e+02
440	6	0.7	403	3	R15628	Mutant Cys(169), Cys	6.28e+02	6	0.7	461	12	R62653	Human protein C.	6.28e+02
441	6	0.7	403	1	R03610	Prepro-alkaline prote	6.28e+02	6	0.7	461	3	R13074	Protein C precursor.	6.28e+02
442	6	0.7	403	4	R20463	Modified alkali prote	6.28e+02	6	0.7	461	3	R13622	Human protein C.	6.28e+02
443	6	0.7	403	4	R20464	Modified alkali prote	6.28e+02	6	0.7	462	26	W33420	Human semenogelin I.	6.28e+02
444	6	0.7	403	1	R04601	Encodes alkaline prot	6.28e+02	6	0.7	462	17	R95853	Human semenogelin I.	6.28e+02
445	6	0.7	404	11	R83232	Epstein-Barr virus ea	6.28e+02	6	0.7	467	9	R48950	Sequence encoded by a	6.28e+02
446	6	0.7	405	12	R62515	Bovine beta3 adrenery	6.28e+02	6	0.7	467	17	R89137	Bombyx mori full leng	6.28e+02
447	6	0.7	406	22	W20788	H. pylori cytoplasmic	6.28e+02	6	0.7	467	18	R96215	Recombinant human D4	6.28e+02
448	6	0.7	406	29	W40035	Novel human protein d	6.28e+02	6	0.7	470	24	W25766	Human CART1.	6.28e+02
449	6	0.7	407	38	W73420	Human secreted protei	6.28e+02	6	0.7	475	21	W11641	Human anti-RSV monocl	6.28e+02
450	6	0.7	407	13	R65494	Marek's disease virus	6.28e+02	6	0.7	477	2	P81868	Sequence of Mycobacte	6.28e+02
451	6	0.7	408	20	W03716	Human autoantigen La	6.28e+02	6	0.7	478	22	W18160	Mycobacterium tubercu	6.28e+02
452	6	0.7	409	22	W17927	Abinding endoglycanase	6.28e+02	6	0.7	478	34	W63035	Mycobacterium tubercu	6.28e+02
453	6	0.7	409	20	W01501	50 kD endoglycanase,	6.28e+02	6	0.7	478	32	W37901	Mycobacterium tubercu	6.28e+02
454	6	0.7	411	16	R90934	HCV NS5 domain antige	6.28e+02	6	0.7	480	38	W74814	Human secreted protei	6.28e+02

528	6	0.7	480 21	W14499	Bacillus amyloliquefa	6.28e+02	601	546 29	W40115	Human alpha-5(IV) col	6.28e+02
529	6	0.7	480 14	R78268	Bacillus amyloliquefa	6.28e+02	502	546 29	W39788	Tobacco PABF protein.	6.28e+02
530	6	0.7	481 9	R72448	Recombinant protein p	6.28e+02	603	551 30	W58884	Human CD33-like prote	6.28e+02
531	6	0.7	481 14	R72448	Bacillus amyloliquefa	6.28e+02	604	555 25	W28503	Timothy grass pollen	6.28e+02
532	6	0.7	483 28	W1405	Bacillus amyloliquefa	6.28e+02	605	556 25	W28502	Timothy grass pollen	6.28e+02
533	6	0.7	484 39	W39274	Granulocytic Ehrlich	6.28e+02	606	560 38	W81965	Human myl-1 protein.	6.28e+02
534	6	0.7	493 39	W3125	PrTIR45 Arginine spe	6.28e+02	607	560 5	R27535	myl protein.	6.28e+02
535	6	0.7	493 4	R20817	Epithelial CD44 Antig	6.28e+02	608	561 1	P31588	Rhopty membrane anti	6.28e+02
536	6	0.7	493 18	R31445	Human epithelial CD4	6.28e+02	609	564 34	W62654	C3 binding protein fr	6.28e+02
537	6	0.7	493 4	P40065	Sequence of human inm	6.28e+02	610	566 31	W59136	Human cdc25B protein.	6.28e+02
538	6	0.7	495 28	W39747	Human marcosR protein	6.28e+02	611	566 32	W37987	Amino acid sequence f	6.28e+02
539	6	0.7	498 3	R33916	pseudorabies virus gx	6.28e+02	612	566 7	R37494	cdc25B.	6.28e+02
540	6	0.7	498 3	R33713	PRV glycoprotein gx	6.28e+02	613	566 13	R66609	Human cdc25B gene pro	6.28e+02
541	6	0.7	499 38	W67010	HCV non-structural pr	6.28e+02	614	566 18	R38213	Human CDC25B phosphat	6.28e+02
542	6	0.7	501 36	W22104	HSV-2 strain SB5 cont	6.28e+02	615	566 28	W35317	Human cdc25B cell cyc	6.28e+02
543	6	0.7	503 34	W21453	Saccharomyces cerevis	6.28e+02	616	567 36	W57770	Human oxidoreductase	6.28e+02
544	6	0.7	504 38	W80513	Hepatitis C virus ant	6.28e+02	617	567 32	W51244	Human calcitonin rece	6.28e+02
545	6	0.7	504 38	W70825	Hepatitis C virus ant	6.28e+02	618	568 32	P80216	Sequence of Mycobacte	6.28e+02
546	6	0.7	504 12	R62659	Hepatitis Type C viru	6.28e+02	619	568 1	R04240	Gamma-glutamyl transp	6.28e+02
547	6	0.7	504 12	R71619	Porcine TrkC NCL isof	6.28e+02	620	575 15	R82214	Bovine herpesvirus-1	6.28e+02
548	6	0.7	504 3	R4553	Chimpanzee HCV clone	6.28e+02	621	575 6	R30045	gE gene deletion from	6.28e+02
549	6	0.7	504 17	R81941	Hepatitis C virus ant	6.28e+02	622	575 9	R47175	Sequence of polypepti	6.28e+02
550	6	0.7	505 24	W10728	Mouse tub gene produc	6.28e+02	623	580 22	W10660	Endoglycoceramidase a	6.28e+02
551	6	0.7	505 27	W36486	Mouse TUB Form II.	6.28e+02	624	584 17	R81626	Human trkC receptor p	6.28e+02
552	6	0.7	505 32	W54367	Mouse tub polypeptide	6.28e+02	625	590 3	R12594	Antigenic portion of	6.28e+02
553	6	0.7	509 3	R33083	PAP-I-protein C fusio	6.28e+02	626	591 9	R49856	Sequence of the env p	6.28e+02
554	6	0.7	510 7	R34674	Insulinoma-associated	6.28e+02	627	591 22	W01679	NS5B protein (residue	6.28e+02
555	6	0.7	510 35	W39358	Pyridine nucleotide t	6.28e+02	628	592 39	W82661	A. thaliana L-galacto	6.28e+02
556	6	0.7	514 2	R10379	Wild type alpha-amyli	6.28e+02	629	593 21	W14794	Human vesicle transpo	6.28e+02
557	6	0.7	514 31	W47004	Glutathione-S-transfe	6.28e+02	630	593 21	W14789	Recombinant modified	6.28e+02
558	6	0.7	514 2	R10378	Mutant alpha-amyliase	6.28e+02	631	593 21	W14789	Hepatitis C virus RNA	6.28e+02
559	6	0.7	514 32	W60749	59 kDa bovine CAM-PDE	6.28e+02	632	596 8	R41433	PT-NANBH virus NS5 pr	6.28e+02
560	6	0.7	514 22	W80337	Bovine lung 59 kDa Ca	6.28e+02	633	597 9	R49857	Sequence of the env p	6.28e+02
561	6	0.7	514 5	R28401	Bovine lung 59 kD CAM	6.28e+02	634	600 39	W82660	Sequence of nuclear r	6.28e+02
562	6	0.7	514 13	R69715	Cyclic-GMP stimulated	6.28e+02	635	604 31	W37163	Cauliflower L-galacto	6.28e+02
563	6	0.7	514 20	W11240	Calcium/calmodulin de	6.28e+02	636	609 37	W33215	Aspergillus oryzae al	6.28e+02
564	6	0.7	514 35	W12421	59 kDa CAM-PDE from b	6.28e+02	637	609 4	R23174	Human h-NUMB-R.	6.28e+02
565	6	0.7	514 34	W77037	Bovine lung Ca2+/calm	6.28e+02	638	612 21	W14790	Mutant thermostable D	6.28e+02
566	6	0.7	520 28	W39748	Human marcosR protein	6.28e+02	639	612 21	W14793	Recombinant modified	6.28e+02
567	6	0.7	520 28	W39743	B. amyloliquefaciens	6.28e+02	640	612 21	W14793	Modified HCV RNA-depe	6.28e+02
568	6	0.7	523 36	W78915	Bovine butyrophilin p	6.28e+02	641	612 21	W14791	Modified HCV RNA-depe	6.28e+02
569	6	0.7	523 37	W73045	Pirv virus G glycopro	6.28e+02	642	612 21	W14792	Modified HCV RNA-depe	6.28e+02
570	6	0.7	529 5	R24948	Sequence encoded by h	6.28e+02	643	615 13	R77276	ORC3 subunit of yeast	6.28e+02
571	6	0.7	529 37	W73046	Chandipura virus G gl	6.28e+02	644	615 2	R10100	E. tenella protein Efi	6.28e+02
572	6	0.7	529 34	W49093	Human wild-type heat	6.28e+02	645	615 23	W22226	S. cerevisiae origin	6.28e+02
573	6	0.7	529 3	R13503	HSF.	6.28e+02	646	616 13	R73005	Aminoacylase O12 cl	6.28e+02
574	6	0.7	530 13	R69711	Cyclic-GMP stimulated	6.28e+02	647	616 13	R73007	Aminoacylase O12 cl	6.28e+02
575	6	0.7	530 22	W18036	Bovine brain 61 kDa C	6.28e+02	648	617 15	R79036	Infectious bovine rhi	6.28e+02
576	6	0.7	530 20	W11243	61 kD brain calcium/c	6.28e+02	649	617 6	R31956	Sequence encoded by t	6.28e+02
577	6	0.7	530 34	W77036	Bovine brain Ca2+/cal	6.28e+02	650	619 39	W89271	Granulocytic Ehrlich	6.28e+02
578	6	0.7	530 35	W12220	61 kDa CAM-PDE DNA en	6.28e+02	651	619 16	R86911	Pneumococcal surfac	6.28e+02
579	6	0.7	530 5	R28395	Bovine brain CAM PDE	6.28e+02	652	619 16	R87598	Pneumococcal surfac	6.28e+02
580	6	0.7	530 32	W60747	61 kDa bovine CAM-PDE	6.28e+02	653	619 12	R63437	Pneumococcal surfac	6.28e+02
581	6	0.7	532 23	W22219	Protein encoded by pv	6.28e+02	654	622 10	R53992	Megakaryocyte potenti	6.28e+02
582	6	0.7	532 16	R31035	Recombinant V8 protea	6.28e+02	655	628 26	W26674	Human CAK1 antigen (m	6.28e+02
583	6	0.7	532 29	W40114	Human alpha-1(IV) col	6.28e+02	656	631 27	W23036	Human cytomagalovirus	6.28e+02
584	6	0.7	534 15	R80164	Mouse signal transduc	6.28e+02	657	631 28	W46444	Cbpa, a choline bindi	6.28e+02
585	6	0.7	535 13	R69731	Cyclic-GMP stimulated	6.28e+02	658	635 14	R72350	Mature human proteint	6.28e+02
586	6	0.7	535 34	W71227	Human 61 kDa CAM-PDE	6.28e+02	659	635 14	R31875	Vitamin K-dependent p	6.28e+02
587	6	0.7	535 32	W60755	Amino acid sequence o	6.28e+02	660	641 33	W61217	Streptococcus pneumon	6.28e+02
588	6	0.7	535 16	R69600	Human GRB-7.	6.28e+02	661	648 39	W87879	A pneumococcal surfac	6.28e+02
589	6	0.7	535 20	W11235	Hippocampus calcium/c	6.28e+02	662	648 33	W62274	Streptococcus pneumon	6.28e+02
590	6	0.7	535 22	W18039	Human brain 61 kDa Ca	6.28e+02	663	648 35	W70336	Streptococcus surfac	6.28e+02
591	6	0.7	535 5	R28411	Protein encoded by cD	6.28e+02	664	650 1	P81137	Pneumococcal surfac	6.28e+02
592	6	0.7	535 35	W77043	Human Ca2+/calmodulin	6.28e+02	665	652 36	W80620	Human protein S	6.28e+02
593	6	0.7	537 36	W49866	Bankia gouldi endoglu	6.28e+02	666	653 5	R27150	S. pneumoniae putativ	6.28e+02
594	6	0.7	537 23	W22220	Protein encoded by pv	6.28e+02	667	657 6	R32657	PSA fragment.	6.28e+02
595	6	0.7	537 26	W34562	Bankia gouldi endoglu	6.28e+02	668	659 15	R89265	PSI protein from C.me	6.28e+02
596	6	0.7	541 6	R33596	HCV CKS-NS5F recombin	6.28e+02	669	673 18	W04208	Heparinase-III.	6.28e+02
597	6	0.7	541 4	R21567	HCV CKS-NS5F - pHCV-4	6.28e+02	670	674 37	W72748	Streptavidin/lucifera	6.28e+02
598	6	0.7	541 7	R33634	HCV CKS-NS5F fusion a	6.28e+02	671	675 13	R70728	Human P-dlg protein.	6.28e+02
599	6	0.7	541 6	R33576	HCV CKS-NS5F recombin	6.28e+02	672	676 3	P70083	Human protein S.	6.28e+02
600	6	0.7	543 38	W55610	Secreted protein clon	6.28e+02	673	676 28	W46454	Human recombinant pro	6.28e+02

674	6	0.7	682 11	R49144	Product of alternativ	6.28e+02	747	6	0.7	858 28	W23314	Human prostate protei	6.28e+02
675	6	0.7	682 17	R87154	Alternatively spliced	6.28e+02	748	6	0.7	870 6	R30729	p100 protein from hum	6.28e+02
676	6	0.7	684 14	R73912	Streptococcus pneumon	6.28e+02	749	6	0.7	873 14	R79459	Infectious laryngotra	6.28e+02
677	6	0.7	684 22	W15287	Human alpha3(IX) coll	6.28e+02	750	6	0.7	877 14	R83657	Human mGluR3.	6.28e+02
678	6	0.7	686 19	R97280	Helicobacter-specific	6.28e+02	751	6	0.7	879 12	R64252	Human mGluR3.	6.28e+02
679	6	0.7	689 17	R88648	Neisseria meningitidi	6.28e+02	752	6	0.7	880 39	W90088	Human G-protein regul	6.28e+02
680	6	0.7	690 4	R23173	Mutant thermotable D	6.28e+02	753	6	0.7	880 31	R23949	Human phosphoinositid	6.28e+02
681	6	0.7	691 7	R34404	Sequence of low molec	6.28e+02	754	6	0.7	886 39	W89377	Staphylococcus aureus	6.28e+02
682	6	0.7	691 7	R34447	N.meningitidis 2169 T	6.28e+02	755	6	0.7	892 4	R23122	Taf DNA polymerase I.	6.28e+02
683	6	0.7	691 7	R88647	Neisseria meningitidi	6.28e+02	756	6	0.7	892 4	R23169	Mutant thermotable D	6.28e+02
684	6	0.7	700 1	R06366	Human alpha-1 collage	6.28e+02	757	6	0.7	902 19	R02250	Human transcription f	6.28e+02
685	6	0.7	709 32	W58988	Homo sapiens fetal ki	6.28e+02	758	6	0.7	904 17	R87147	Protocadherin clone 4	6.28e+02
686	6	0.7	711 17	R88643	Neisseria meningitidi	6.28e+02	759	6	0.7	904 11	R58907	Human protocadherin-4	6.28e+02
687	6	0.7	711 26	W14619	HTR Tbp2 protein from	6.28e+02	760	6	0.7	921 35	W72033	HSV-2 strain S85 Cont	6.28e+02
688	6	0.7	711 9	R48222	N.meningitidis IM2169	6.28e+02	761	6	0.7	929 24	W14593	Streptococcus pneumon	6.28e+02
689	6	0.7	714 9	R47244	Fragment of 101 kd pr	6.28e+02	762	6	0.7	930 8	R41431	PT-NANBH virus BHC-11	6.28e+02
690	6	0.7	715 23	W06593	Amino acid sequence o	6.28e+02	763	6	0.7	931 15	R79009	Human DNA repair prot	6.28e+02
691	6	0.7	716 1	P95783	L-phenylalanine ammon	6.28e+02	764	6	0.7	934 5	P20016	Sequence of p20, VP4	6.28e+02
692	6	0.7	716 1	P81099	Sequence of Rhodospir	6.28e+02	765	6	0.7	934 20	W07637	P. suis leukotoxin ge	6.28e+02
693	6	0.7	716 1	P83141	Sequence of L-phenyla	6.28e+02	766	6	0.7	949 33	W61349	Human telomerase prot	6.28e+02
694	6	0.7	716 13	R70682	Stabilised phenylalan	6.28e+02	767	6	0.7	952 35	W63118	Human adenovirus 5 he	6.28e+02
695	6	0.7	716 1	P80513	L-phenylalanine ammon	6.28e+02	768	6	0.7	952 36	W79539	Adenovirus serotype 5	6.28e+02
696	6	0.7	716 28	W38192	Maize ZCARECROW ZCR p	6.28e+02	769	6	0.7	966 23	W19916	Drosophila melanogast	6.28e+02
697	6	0.7	728 5	R27149	Mouse trkC proto-onco	6.28e+02	770	6	0.7	968 36	W79538	Adenovirus serotype 2	6.28e+02
698	6	0.7	728 12	R62022	Murine trkC protein.	6.28e+02	771	6	0.7	974 32	W55960	Human transient recep	6.28e+02
699	6	0.7	728 6	R30884	Partial trkC gene pro	6.28e+02	772	6	0.7	985 21	W15191	Aspergillus oryzae al	6.28e+02
700	6	0.7	729 17	R89275	Yeast coagulation pro	6.28e+02	773	6	0.7	987 31	W57057	Class II trans activa	6.28e+02
701	6	0.7	732 2	R10426	Human testicular angi	6.28e+02	774	6	0.7	991 14	R77313	Porphyromonas gingiva	6.28e+02
702	6	0.7	734 23	W13657	Streptomyces viridosp	6.28e+02	775	6	0.7	996 28	W33624	Eimeria tenella 45 kd	6.28e+02
703	6	0.7	736 3	R13949	SUP-B27 t(1;19) trans	6.28e+02	776	6	0.7	998 4	P93706	Sequence of the anti b	6.28e+02
704	6	0.7	737 13	R70186	Arg-gingipain-1.	6.28e+02	777	6	0.7	998 28	W33621	Eimeria tenella 45 kd	6.28e+02
705	6	0.7	737 29	W34946	New DNA sequence iso	6.28e+02	778	6	0.7	1015 35	W63675	Polypeptide having ag	6.28e+02
706	6	0.7	738 31	W56163	Arg-gingipain high mo	6.28e+02	779	6	0.7	1021 24	W23281	Alloreactive associat	6.28e+02
707	6	0.7	739 12	R17618	Murine trkC K3 isofor	6.28e+02	780	6	0.7	1045 31	W57448	M. tuberculosis isole	6.28e+02
708	6	0.7	742 3	R15358	E2A/pr1 fusion protei	6.28e+02	781	6	0.7	1046 19	W02156	Periplasmic chitodext	6.28e+02
709	6	0.7	745 1	P90615	Polypeptide with isoa	6.28e+02	782	6	0.7	1055 22	W14319	Thermotoga neopolitan	6.28e+02
710	6	0.7	750 32	W62186	Human RAD54 N-termina	6.28e+02	783	6	0.7	1078 13	R71704	Collagen alpha 1 (III	6.28e+02
711	6	0.7	754 4	R23172	Mutant thermotable D	6.28e+02	784	6	0.7	1079 37	W81504	Short form of TPR mol	6.28e+02
712	6	0.7	784 30	W53570	Cucumber raffinose sy	6.28e+02	785	6	0.7	1103 8	R39631	Neurofibromatosis typ	6.28e+02
713	6	0.7	785 36	W72195	HSV-2 strain S85 Cont	6.28e+02	786	6	0.7	1106 37	W80312	CIITA (class II trans	6.28e+02
714	6	0.7	792 11	R56502	X-prolyl-di-peptidyl-a	6.28e+02	787	6	0.7	1124 8	R41439	PT-NANBH NS5-NS3-core	6.28e+02
715	6	0.7	796 20	W08087	Infective bursa prote	6.28e+02	788	6	0.7	1130 14	R74452	Class II transactivat	6.28e+02
716	6	0.7	797 38	W81963	Human myl/RAR-alpha f	6.28e+02	789	6	0.7	1130 16	R81569	Class II transactivat	6.28e+02
717	6	0.7	797 26	W35577	TNF-R1-DD ligand prot	6.28e+02	790	6	0.7	1130 37	W80313	CIITA (class II trans	6.28e+02
718	6	0.7	797 5	R27533	myl/RAR-alpha fusion.	6.28e+02	791	6	0.7	1130 31	W57056	Class II trans activa	6.28e+02
719	6	0.7	798 7	R33630	HCV CKS-R55 EF fusion	6.28e+02	792	6	0.7	1132 31	W56113	Human telomerase reve	6.28e+02
720	6	0.7	800 4	R23171	Mutant thermotable D	6.28e+02	793	6	0.7	1132 31	W46957	Human telomerase reve	6.28e+02
721	6	0.7	802 8	R45001	Cellulose synthase op	6.28e+02	794	6	0.7	1132 35	W71376	Human telomerase cata	6.28e+02
722	6	0.7	807 31	W46997	Human telomerase reve	6.28e+02	795	6	0.7	1134 29	W46859	Bacillus thuringiens	6.28e+02
723	6	0.7	808 28	W38188	Arabidopsis SCARECROW	6.28e+02	796	6	0.7	1135 30	W33808	Mouse ataxin-2.	6.28e+02
724	6	0.7	819 3	R13668	SUP-B27 t(1;19) trans	6.28e+02	797	6	0.7	1150 31	W47006	Glutathione-S-transfe	6.28e+02
725	6	0.7	823 23	W13668	DhpA-mel chimeric pro	6.28e+02	798	6	0.7	1150 33	W62541	Bacillus thuringiens	6.28e+02
726	6	0.7	823 19	W00364	Human CDC27.	6.28e+02	799	6	0.7	1150 29	W46858	Bacillus thuringiens	6.28e+02
727	6	0.7	824 14	R75848	H-NUC retinoblastoma	6.28e+02	800	6	0.7	1150 33	W61350	Human telomerase prot	6.28e+02
728	6	0.7	825 5	R27148	Adult porcine trkC pr	6.28e+02	801	6	0.7	1156 29	W46857	Bacillus thuringiens	6.28e+02
729	6	0.7	825 17	R81631	Human trkC receptor p	6.28e+02	802	6	0.7	1156 29	W46856	Bacillus thuringiens	6.28e+02
730	6	0.7	825 13	R13951	E2A/pr1 fusion protei	6.28e+02	803	6	0.7	1157 9	R48678	Insecticidal protoxin	6.28e+02
731	6	0.7	825 12	R62021	Porcine trkC protein.	6.28e+02	804	6	0.7	1169 18	R96126	Bacillus thuringiens	6.28e+02
732	6	0.7	825 4	R22337	Sequence of the "gb"	6.28e+02	805	6	0.7	1189 31	W47008	Glutathione-S-transfe	6.28e+02
733	6	0.7	825 6	R30883	trkC gene prod.	6.28e+02	806	6	0.7	1199 31	W47007	Glutathione-S-transfe	6.28e+02
734	6	0.7	830 37	W80314	CIITA (class II trans	6.28e+02	807	6	0.7	1207 37	W80311	CIITA (class II trans	6.28e+02
735	6	0.7	830 17	R81627	Human trkC receptor p	6.28e+02	808	6	0.7	1250 3	R12599	Portion of PT-NANBH v	6.28e+02
736	6	0.7	835 17	R96206	Invasin protein.	6.28e+02	809	6	0.7	1266 31	W48721	Human cytoplasmic iso	6.28e+02
737	6	0.7	836 17	R87153	Alternatively spliced	6.28e+02	810	6	0.7	1285 31	W47000	HIS tagged thioedoxi	6.28e+02
738	6	0.7	836 11	R58912	Product of alternativ	6.28e+02	811	6	0.7	1302 15	R84741	Multidrug resistance	6.28e+02
739	6	0.7	837 39	W75425	Human aggrecan degrad	6.28e+02	812	6	0.7	1305 37	W79274	Protein kinase GAK.	6.28e+02
740	6	0.7	839 17	R81625	Human trkC receptor p	6.28e+02	813	6	0.7	1306 1	R04111	Human angiotensin con	6.28e+02
741	6	0.7	839 12	R71621	Porcine trkC K2 isofo	6.28e+02	814	6	0.7	1306 34	W68155	Human angiotensin con	6.28e+02
742	6	0.7	847 39	W85603	Hexosaminidase enzyme	6.28e+02	815	6	0.7	1311 37	W79275	Protein kinase GAK.	6.28e+02
743	6	0.7	847 6	R33419	CD22-beta.	6.28e+02	816	6	0.7	1312 30	W33807	Human ataxin-2.	6.28e+02
744	6	0.7	850 21	W11942	gd.trkC fusion usefu	6.28e+02	817	6	0.7	1313 33	W60213	Spinocerebellar ataxi	6.28e+02
745	6	0.7	856 4	R23170	Mutant thermotable D	6.28e+02	818	6	0.7	1323 10	R55248	N-methyl-D-aspartic a	6.28e+02
746	6	0.7	858 9	R47242	101 kd protein of hum	6.28e+02	819	6	0.7	1338 8	R41731	High molecular weight	6.28e+02

820	6	0.7	1405 31	W56101	Enhanced green fluore	6.28e+02	893	W40038	HCV polyprotein.	6.28e+02
821	6	0.7	1422 16	R82059	Hepatitis GB virus (H	6.28e+02	894	R40119	HCV genomic amino aci	6.28e+02
822	6	0.7	1427 2	R10584	Human 160kD mediator	6.28e+02	895	R31621	Hepatitis C virus (HC	6.28e+02
823	6	0.7	1451 5	R27819	CCVinsavc spike prote	6.28e+02	896	R54099	NANBHV E1/E2 protein.	6.28e+02
824	6	0.7	1475 27	R08221	Recombinant alpha amy	6.28e+02	897	R35207	Hepatitis C virus pro	6.28e+02
825	6	0.7	1477 27	W30294	Non-typeable Haemophi	6.28e+02	898	R36178	Murine Ataxia-telangi	6.28e+02
826	6	0.7	1477 12	R63506	Haemophilus high mole	6.28e+02	899	R31103	Merosin major subunit	6.28e+02
827	6	0.7	1477 8	R41724	High molecular weight	6.28e+02	900	R72204	HSV-2 strain SB5 Cont	6.28e+02
828	6	0.7	1477 8	R41728	High molecular weight	6.28e+02	901	R31950	Human bg protein asso	6.28e+02
829	6	0.7	1498 14	R77084	Rat sulphophylurea rec	6.28e+02	902	R10145	Cephalosporin antibio	6.28e+02
830	6	0.7	1498 14	R77085	Hamster sulphophylurea	6.28e+02	903	R31949	Human bg protein asso	6.28e+02
831	6	0.7	1529 8	R41732	High molecular weight	6.28e+02	904	R33396	Human PKD1 polypeptid	6.28e+02
832	6	0.7	1531 25	W23597	Human Lxsn1 short iso	6.28e+02	905	R33830	Human PKD1 polypeptid	6.28e+02
833	6	0.7	1547 1	R07304	IgA1 protease.	6.28e+02	906	R400870	Polycystic kidney dis	6.28e+02
834	6	0.7	1571 16	R91047	Alpha-D-glucosyltrans	6.28e+02	907	R30302	Polycystic kidney dis	6.28e+02
835	6	0.7	1582 14	R77087	Rat sulphophylurea rec	6.28e+02	908	R75916	Polycystic kidney dis	6.28e+02
836	6	0.7	1582 14	R77088	Hamster sulphophylurea	6.28e+02	909	R87539	Polycystic kidney dis	6.28e+02
837	6	0.7	1598 27	W30291	Non-typeable Haemophi	6.28e+02	910	R47861	Alpha 2-Macroglobulin	6.28e+02
838	6	0.7	1601 27	W30292	Non-typeable Haemophi	6.28e+02	911	R60517	Human alpha-2-MR.	6.28e+02
839	6	0.7	1612 32	W65088	R. prowazekii S-layer	6.28e+02	912	R60517	Hybrid smg/tylg ORF1	6.28e+02
840	6	0.7	1684 24	W46761	Amino acid sequence o	6.28e+02	913	W22611	Platenolide synthase	6.28e+02
841	6	0.7	1684 24	W25671	hABC3 protein.	6.28e+02	914	W23716	Platenolide synthase	6.28e+02
842	6	0.7	1687 17	R96033	P. gingivalis haemagg	6.28e+02	915	W22606	Platenolide synthase	6.28e+02
843	6	0.7	1687 36	W69495	Haemagglutinin protei	6.28e+02	916	R97210	Human kidney calcium	6.28e+02
844	6	0.7	1704 13	R70188	Arg-gingipain-2 prepo	6.28e+02	917	R97211	Human calcium sensor	6.28e+02
845	6	0.7	1704 29	W46771	Amino acid sequence o	6.28e+02	918	W43311	Human calcium sensor	6.28e+02
846	6	0.7	1704 28	W34843	Arg-gingipain high mo	6.28e+02	919	W43312	Human kidney calcium	6.28e+02
847	6	0.7	1706 24	W24796	PrTR antigenic protei	6.28e+02	920	W43314	Human placental calci	6.28e+02
848	6	0.7	1723 19	W00645	Mouse DEC-205.	6.28e+02	921	W43314	Human parathyroid cal	6.28e+02
849	6	0.7	1766 1	R92041	Sequence encoded in t	6.28e+02	922	R97209	Human placental calci	6.28e+02
850	6	0.7	1786 1	R90158	Protein sequence of h	6.28e+02	923	W52846	A. mediterranei rifam	6.28e+02
851	6	0.7	1810 16	R94563	Chicken cytotoxicin.	6.28e+02	924	R44929	T. niveum Cyclosporin	6.28e+02
852	6	0.7	1886 31	W54241	Rattus norvegicus mut	6.28e+02	925	W15063	Bicyclic neurotrophin	5.46e+03
853	6	0.7	1896 36	W72095	HSV-2 strain SB5 Cont	6.28e+02	926	W23510	Purified cis-9,10-oct	5.46e+03
854	6	0.7	1996 26	W22607	Platenolide synthase	6.28e+02	927	W06237	ATM epitope #3.	5.46e+03
855	6	0.7	1996 26	W23717	Platenolide synthase	6.28e+02	928	R07018	Residues 301-324 of H	5.46e+03
856	6	0.7	2001 25	W23596	Human Lxsn1 longer is	6.28e+02	929	R67649	RNA-ase-based catalys	5.46e+03
857	6	0.7	2115 32	W59276	Rubella virus RA27/3	6.28e+02	930	W32240	B chain sequence of a	5.46e+03
858	6	0.7	2201 22	W01680	HCV NS2-NS5B non-stru	6.28e+02	931	W05247	Glutamate-gated chlor	5.46e+03
859	6	0.7	2205 15	R79048	Infectious rubella vi	6.28e+02	932	W24630	H. pylori transmembra	5.46e+03
860	6	0.7	2247 25	W27126	Equine rhinovirus 1 (	6.28e+02	933	P61103	N-terminally truncate	5.46e+03
861	6	0.7	2261 1	R90164	Peptide encoded by co	6.28e+02	934	W02150	Anti-obesity protein.	5.46e+03
862	6	0.7	2273 37	W70398	ATP binding cassette	6.28e+02	935	W24511	Anti-Factor-IX humani	5.46e+03
863	6	0.7	2289 25	W14987	Protein derived from	6.28e+02	936	W10540	Humanised murine anti	5.46e+03
864	6	0.7	2301 1	R92047	Sequence encoded in t	6.28e+02	937	R05948	HNK-20 variable heavy	5.46e+03
865	6	0.7	2354 8	R41435	PT-NANBH virus non-st	6.28e+02	938	W10665	Human TGF-beta-like c	5.46e+03
866	6	0.7	2435 5	R25135	HCV polypeptide 1.	6.28e+02	939	P61103	Hepatitis B virus pol	5.46e+03
867	6	0.7	2436 5	R28582	HCV amino acid sequen	6.28e+02	940	R71243	Human erythropoietin	5.46e+03
868	6	0.7	2436 1	R92050	Sequence encoded in t	6.28e+02	941	W20609	H. pylori secreted or	5.46e+03
869	6	0.7	2462 1	R90288	Peptide encoded by co	6.28e+02	942	R95987	Nucleic acid recognit	5.46e+03
870	6	0.7	2485 11	R59921	RAS associated GAP NF	6.28e+02	943	P80436	Recombinant human glu	5.46e+03
871	6	0.7	2485 11	R59922	RAS associated GAP NF	6.28e+02	944	W15933	Antibody 7G12 light c	5.46e+03
872	6	0.7	2707 24	W27161	Mouse receptor ME2.	6.28e+02	945	R05081	MG-1 antigen.	5.46e+03
873	6	0.7	2772 2	R08123	Hepatitis C virus pol	6.28e+02	946	W10079	N-delta-3/C-delta-249	5.46e+03
874	6	0.7	2816 7	R34009	HCV-1 polyprotein.	6.28e+02	947	R05091	Vector PSP 35K-1A-10	5.46e+03
875	6	0.7	2818 21	W12268	Human neurofibromin.	6.28e+02	948	R55890	G6F2BPP3.	5.46e+03
876	6	0.7	2818 4	R23280	Nfl gene product.	6.28e+02	949	W14354	Human DNase I variant	5.46e+03
877	6	0.7	2894 5	R24440	Composite HCV HC-J1/C	6.28e+02	950	W14355	Human DNase I variant	5.46e+03
878	6	0.7	2894 13	R70230	Composite hepatitis C	6.28e+02	951	W21718	RIP fusion protein, R	5.46e+03
879	6	0.7	2955 2	R08124	Hepatitis C virus put	6.28e+02	952	W14853	Subtilisin DV variant	5.46e+03
880	6	0.7	3010 4	R20091	Non-A, non-B viral ge	6.28e+02	953	R07140	H.influenzae (HI) e"	5.46e+03
881	6	0.7	3010 4	R20091	Non-A, non-B viral ge	6.28e+02	954	R55656	G6F2BPP2.	5.46e+03
882	6	0.7	3010 10	R53417	Blood transmissible N	6.28e+02	955	W32430	Mycobacterium tubercu	5.46e+03
883	6	0.7	3011 36	W77397	Hepatitis C virus H77	6.28e+02	956	W24574	gp(1-263)-V3 fusion g	5.46e+03
884	6	0.7	3011 16	R90931	Hepatitis C virus pol	6.28e+02	957	R78611	Murine Fas antigen ex	5.46e+03
885	6	0.7	3011 36	W77398	Hepatitis C virus-H C	6.28e+02	958	R64769	HSV-2 ISP35 protein.	5.46e+03
886	6	0.7	3011 26	W34480	HCV polyprotein.	6.28e+02	959	W06319	Human mitogen-activat	5.46e+03
887	6	0.7	3011 8	R40120	HCV genomic amino aci	6.28e+02	960	P91388	Glutamine synthetase	5.46e+03
888	6	0.7	3011 13	R67568	Hepatitis C virus gen	6.28e+02	961	W29148	Polylactonase inh	5.46e+03
889	6	0.7	3011 14	R79232	HCV sequence.	6.28e+02	962	W11820	Murine myeloid-lineag	5.46e+03
890	6	0.7	3011 12	R66995	Hepatitis C virus gen	6.28e+02	963	R13515	P.denitrificans COB W	5.46e+03
891	6	0.7	3011 4	R22154	NANBH Hutch c59 isola	6.28e+02	964	R77099	Myxoma virus SEPP-1	5.46e+03
892	6	0.7	3011 4	R21519	Completed HCV sequence	6.28e+02	965	W02577	Recombinant HIV-2 tra	5.46e+03







Tolerisation: self-epitope; antigen; autoimmune disease;  
 KW autoantigen; HLA; human leukocyte antigen; T-cell; thymocyte;  
 KW pemphigus vulgaris; desmoglein; multiple sclerosis;  
 KW herpes simplex virus; adenovirus; phosphomannomutase;  
 KW human papillomavirus; Epstein-Barr virus; DNA polymerase;  
 KW influenza; haemagglutinin; reovirus; sigma protein.  
 OS Homo sapiens.  
 PN WO9627387-A1.  
 PD 12-SEP-1996.  
 PF 07-MAR-1996; U03182.  
 PR 07-MAR-1996; US-400796.  
 PA (HARD ) HARVARD COLLEGE.  
 PI Strominger JL, Wucherpfennig KW;  
 DR WPI: 96-425218/42.  
 PT pemphigus vulgaris auto-antigens and multiple sclerosis non-self  
 PT antigens - useful in disease treatment, and method for  
 PT identification of other self and non-self antigens implicated in  
 PT auto-immune disease  
 PS Claim 1: Page 38; 58pp; English.  
 CC Pharmaceutical preparations for tolerisation to antigens comprise  
 CC either an isolated human non-collagen or non-mysin basic protein  
 CC (MBP) polypeptide which is capable of tolerising an individual to an  
 CC autoantigen; or an isolated human pathogen polypeptide capable of  
 CC tolerising an individual to that polypeptide. In both cases, the  
 CC polypeptide (whether self or non-self) includes an amino acid  
 CC sequence corresponding to a sequence motif for a MHC class II  
 CC protein, such as HLA-DR, which is associated with a human autoimmune  
 CC disease, and which binds to the polypeptide to activate autoreactive  
 CC T-cells in individuals with the autoimmune disease. This peptide is  
 CC derived from the human desmoglein 3 protein (amino acids 97-111)  
 CC and is implicated as a self epitope in pemphigus vulgaris. Peptides  
 CC derived from the human desmoglein protein are described in W04841-47.  
 SQ Sequence 15 AA;  
 Query Match 0.8%; Score 7; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred.No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 4 fvdvknk 10  
 |||||  
 QY 542 FVVDKNT 548  
 RESULT 4  
 ID R93961 standard; peptide; 19 AA.  
 AC R93961:  
 DT 05-JUL-1996 (first entry)  
 DE Peptide combining with anti-interdermal cellular antibody.  
 KW anti-interdermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 129556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 96-045392/05.  
 PT Anti-interdermal cellular antibody-combining peptide - which can  
 PT be immobilised on column to form adsorbent useful for treating  
 PT diseases related to the antibody  
 PS Example 1; Page 4; 7pp; Japanese.  
 CC New peptides are disclosed which contain at least 5 contiguous amino  
 CC acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe  
 CC Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg  
 CC Glu Glu (see R93960), the peptide not containing more than 50 residues.  
 CC The peptide combines with anti-interdermal cellular antibody. It can  
 CC be immobilised on a carrier to prepare an adsorbent useful for the  
 CC treatment of diseases related to anti-interdermal cellular antibody.  
 CC The present sequence is a specific example of the new peptides.  
 SQ Sequence 19 AA;  
 Query Match 0.8%; Score 7; DB 16; Length 19;  
 Best Local Similarity 100.0%; Pred.No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 fvdvknk 17  
 |||||  
 QY 542 FVVDKNT 548  
 RESULT 5  
 ID R47000 standard; Protein; 21 AA.  
 AC R47000;  
 DT 16-SEP-1994 (first entry)  
 DE LAM Blast-1 with N-acetylglucosamine chain position 88-108.  
 KW Naturally-occurring; immunomodulatory protein; human; therapy; class I;  
 KW major histocompatibility complex; class II; allotype; type I diabetes;  
 KW autoimmune disease; rheumatoid arthritis; T-cell-mediated response;  
 KW multiple sclerosis; transplant rejection; vaccine; MHC.  
 OS Homo sapiens.  
 PN WO9404171-A.  
 PD 03-MAR-1994.  
 PF 11-AUG-1993; U07545.  
 PR 11-AUG-1993; US-925460.  
 PR 15-JUN-1993; US-925460.  
 PA (HARD ) HARVARD COLLEGE.  
 PI Chicz RM, Hedley ML, Stern LJ, Strominger JL, Urban RG;  
 PI Vignali DA;  
 DR WPI: 94-082825/10.  
 PT Novel immunomodulatory peptide(s) and nucleic acids - useful for  
 PT treatment of auto-immune diseases, transplant rejection and for  
 PT vaccination  
 PS Disclosure; Page 48; 139pp; English.  
 CC The sequences given in R49291-505 and R46981-7038 represent peptide  
 CC fragments of naturally-occurring immunomodulatory proteins. These  
 CC fragments are between 10-30 residues in length and bind to a human  
 CC major histocompatibility complex (MHC) class II allotype. These  
 CC peptides may be used for therapy of autoimmune diseases, such as  
 CC type I diabetes, rheumatoid arthritis and multiple sclerosis, and to  
 CC reduce transplant rejection. They may also be used for vaccination  
 CC providing an exclusively T-cell-mediated response, which can be  
 CC class I or class-II based, or both, depending on the length and  
 CC character of the immunogenic peptides.  
 SQ Sequence 21 AA;  
 Query Match 0.8%; Score 7; DB 10; Length 21;  
 Best Local Similarity 100.0%; Pred.No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 3 qsgalyi 9  
 |||||  
 QY 692 QSGALYI 698  
 RESULT 6  
 ID R93962 standard; peptide; 21 AA.  
 AC R93962:  
 DT 05-JUL-1996 (first entry)  
 DE Peptide combining with anti-interdermal cellular antibody.  
 KW anti-interdermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 129556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS ) KURARAY CO LTD.  
 DR WPI: 96-045392/05.  
 PT Anti-interdermal cellular antibody-combining peptide - which can  
 PT be immobilised on column to form adsorbent useful for treating  
 PT diseases related to the antibody  
 PS Example 2; Page 5; 7pp; Japanese.  
 CC New peptides are disclosed which contain at least 5 contiguous amino  
 CC acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe  
 CC Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg  
 CC Glu Glu (see R93960), the peptide not containing more than 50 residues.  
 CC The peptide combines with anti-interdermal cellular antibody. It can  
 CC be immobilised on a carrier to prepare an adsorbent useful for the

CC treatment of diseases related to anti-interepidermal cellular antibody.  
 CC The present sequence is a specific example of the new peptides.  
 SQ Sequence 21 AA;

Query Match 0.8%; Score 7; DB 16; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 fvdvdknt 19  
 |||||  
 QY 542 FVVDKNT 548

RESULT 7

ID R07135 standard; protein; 26 AA.  
 AC R07135;  
 DT 24-JAN-1991 (first entry)  
 DE Peptide with hairpin structure to stabilise NA hybridisation (III).  
 KW Nucleic acid hybridisation; hairpin structure; stabilisation;  
 KW peptide-oligonucleotide conjugate; probe; hinge.  
 OS Synthetic.  
 PN W09010713-A.  
 PD 20-SEP-1990.  
 PF 13-MAR-1990; F00168.  
 PR 14-MAR-1989; FR-003303.  
 PA (OPAL-) OPALE BIOTECHNOLOGY.  
 PI Prieur B;  
 DR WPI: 90-305035/40.  
 PT Stabilisation of nucleic acid hybridisation - using peptide with  
 PT hair-pin structure to link complementary strands  
 PS Example 4; Page 28; 21pp; French.  
 CC Two amino acid sequences are attached by a "hinge" that does not  
 CC interact with DNA or RNA, and comprises 2-4 amino acids selected  
 CC from Gly, Asn or Gln.  
 CC The hybridisation of a first DNA or RNA sequence, having an  
 CC alpha- or beta-anomeric configuration, with a target sequence  
 CC complementary to the first is stabilised by complexing the resulting  
 CC double helix with a peptide like this one, which is capable of  
 CC adopting a hairpin structure between the two strands of the double  
 CC helix. The process is applicable to nucleic acid hybridisation  
 CC probes and to nucleic acid sequences used therapeutically to block  
 CC expression of exogenous (e.g. viral or microbial) or endogenous  
 CC genes.  
 CC The effect of stabilisation by this peptide, was studied by  
 CC elevating the temp. Semi-transition occurred at 48 degrees C  
 CC compared to 32 degrees C for oligonucleotides not conjugated to the  
 CC peptide. Therefore the conjugated form shows a stronger  
 CC hybridisation potential.  
 CC See also R07133-39.  
 SQ Sequence 26 AA;

Query Match 0.8%; Score 7; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 ssggsfv 17  
 |||||  
 QY 732 SSGGSFV 738

RESULT 8

ID R93960 standard; peptide; 30 AA.  
 AC R93960;  
 DT 05-JUL-1996 (first entry)  
 DE Peptide combining with anti-interepidermal cellular antibody.  
 KW anti-interepidermal cellular antibody; autoantibody; adsorbent.  
 OS Synthetic.  
 PN J07309893-A.  
 PD 28-NOV-1995.  
 PF 18-MAY-1994; 129556.  
 PR 18-MAY-1994; JP-129556.  
 PA (KURS) KURARAY CO LTD.  
 DR WPI: 96-045392/05.

PT Anti-interepidermal cellular antibody-combining peptide - which can  
 PT be immobilised on column to form adsorbent useful for treating  
 PT diseases related to the antibody.  
 PS Claim 1; Page 2; 7pp; Japanese.  
 CC A new peptide is disclosed which contains at least 5 contiguous amino  
 CC acids from the sequence Val Gly Ile Asp Gln Pro Phe Gly Ile Phe  
 CC Val Val Asp Lys Asn Thr Gly Asp Ile Asn Ile Thr Ala Ile Val Asp Arg  
 CC Glu Glu (the present sequence), the peptide not containing more than  
 CC 50 residues. The peptide combines with anti-interepidermal cellular  
 CC antibody. It can be immobilised on a carrier to prepare an adsorbent  
 CC useful for the treatment of diseases related to anti-interepidermal  
 CC cellular antibody.  
 SQ Sequence 30 AA;

Query Match 0.8%; Score 7; DB 16; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 fvdvdknt 17  
 |||||  
 QY 542 FVVDKNT 548

RESULT 9

ID W43445 standard; peptide; 37 AA.  
 AC W43445;  
 DT 15-MAY-1998 (first entry)  
 DE Scorpion toxin-related peptide SEQ ID NO:5.  
 KW Scorpion; toxin-related peptide; disulphide bridge; K+ channel;  
 KW potassium channel; toxic.  
 OS Butus martseni.  
 PN J09278797-A.  
 PD 28-OCT-1997.  
 PF 05-APR-1996; 118129.  
 PR 05-APR-1996; JP-118129.  
 PA (SUNK) SUNTORY LTD.  
 DR WPI: 98-022250/03.  
 PT A scorpion toxin-related peptide - allows separation of its toxic  
 PT activity from its activity against potassium channels  
 PS Claim 5; Page 7; 8pp; Japanese.  
 CC The present sequence represents a scorpion toxin-related peptide,  
 CC which has no more than 3 disulphide bridges. The peptide allows  
 CC separation of its activity against K+ channels (potassium channels)  
 CC from its lethal toxicity.  
 SQ Sequence 37 AA;

Query Match 0.8%; Score 7; DB 28; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 kcgggsc 13  
 |||||  
 QY 199 KCQGGSC 205

RESULT 10

ID W42471 standard; protein; 119 AA.  
 AC W42471;  
 DT 22-JUN-1998 (first entry)  
 DE NEW humanised heavy chain variable region.  
 KW Neutralising antibody; monoclonal antibody; MAB; 2B6; NEW; mouse;  
 KW interleukin-5; IL-5; human; eosinophil; asthma; allergic rhinitis;  
 KW atopic dermatitis; therapy; diagnosis; humanised antibody.  
 OS Chimeric - Mus musculus.  
 OS Chimeric - Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Region 1..30  
 FT /label= FRI  
 FT /note= "NEW framework region 1"  
 FT Region 31..35  
 FT /label= CDRI  
 FT /note= "2B6 complementarity determining region 1"  
 FT Region 36..49

FT /label= FR2  
 FT /note= "NEW framework region 2"  
 FT 50..65  
 FT /label= CDR2  
 FT /note= "2B6 complementarity determining region 2"  
 FT 66..97  
 FT /label= FR3  
 FT /note= "NEW framework region 3"  
 FT 98..108  
 FT /label= CDR3  
 FT /note= "2B6 complementarity determining region 3"  
 FT 109..119  
 FT /label= FR4  
 FT /note= "NEW framework region 4"  
 FT W09748418-A1.  
 PN 24-DEC-1997.  
 PD 20-JUN-1997; U10769.  
 PF 21-JUN-1996; US-667769.  
 PR (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA Appelbaum ER, Cook RM;  
 PI WPI: 98-062853/06.  
 DR N-PSDB; V03505.  
 CC Use of neutralising antibody to human interleukin-5 - that does not  
 PT block binding to receptor alpha chain, to treat e.g. asthma,  
 PT allergic rhinitis or atopic dermatitis  
 PS Example 4; Fig 12; 116pp; English.  
 CC This polypeptide comprises a humanised antibody heavy chain  
 CC variable region composed of human NEW framework regions and  
 CC complementarity determining regions (see W42457-59) derived from  
 CC anti-human interleukin-5 (hIL-5) murine monoclonal antibody (MAB)  
 CC 2B6 heavy chain (see W42451). It is encoded by a DNA construct  
 CC (see V03505) in vector pCNIIL5NEW. The invention provides  
 CC antibodies, especially altered, chimeric and humanised antibodies  
 CC which are characterised by hIL-5 specificity, neutralising activity  
 CC and affinity for hIL-5. The antibodies are useful for treating  
 CC hIL-5-mediated disorders such as asthma, allergic rhinitis and  
 CC atopic dermatitis, and can also be used in the diagnosis of such  
 CC conditions by measurement (e.g. by ELISA) of endogenous hIL-5  
 CC levels. Also provided are vectors and transformed host cells for  
 CC expression of the novel antibodies.  
 SQ Sequence 119 AA;

Query Match 0.8%; Score 7; DB 29; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 5.26e-01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 sqtltlt 21  
 |||||  
 QY 572 SQTTLT 578

RESULT 11  
 ID R88109 standard; peptide: 120 AA.  
 AC R88109;  
 DT 25-JUL-1996 (first entry)  
 DE Murine anti-Protein C MAB HPC-4 VH gamma mature peptide.  
 KW Epitope: activation; heavy chain; protein C; vitamin K; plasma protein;  
 KW zymogen; cleavage; mouse; humanised antibody; variable region;  
 KW light chain; inhibition; anticoagulant; coagulation; tumour.  
 OS Mus musculus.  
 PN W09534652-A1.  
 PD 21-DEC-1995.  
 PF 09-JUN-1995; U07372.  
 PR 10-JUN-1994; US-259321.  
 PA (OKLA-) OKLAHOMA MED RES FOUND.  
 PI Esmon CT, Rezaie A;  
 DR WPI: 96-049681/05.  
 CC N-PSDB; T09300.  
 PT Calcium-binding monoclonal antibody immunoreactive with Protein C -  
 PT inhibits Protein C anticoagulant activation by  
 PT thrombin-thrombomodulin, e.g. for treating tumours  
 PS Claim 2; Page 29; 41pp; English.  
 CC This is the amino acid sequence of the mature peptide from the murine

CC anti-protein C monoclonal antibody HPC-4 heavy chain variable region.  
 CC HPC-4 recognises the activation peptide region (R88106) of the heavy  
 CC chain of protein C, a vitamin K-dependent plasma protein zymogen.  
 CC Protein C is converted to activated protein C (APC) by cleavage between  
 CC the Arg-Leu amino acid contained within the activation peptide sequence.  
 CC HPC-4 prevents protein C activation to APC by binding to this region.  
 CC The DNA sequences encoding the variable regions of the heavy and light  
 CC chains of the antibody (T09299-302) were used to construct humanised  
 CC antibodies using the PCR primers T09303-9. The humanised antibodies are  
 CC useful as inhibitors of coagulation and can be used for the treatment of  
 CC tumours by inhibiting the anticoagulant activity of APC by preventing  
 CC conversion of protein C to APC.  
 SQ Sequence 120 AA;

Query Match 0.8%; Score 7; DB 17; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 5.26e-01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 15 sqtltlt 21  
 |||||  
 QY 572 SQTTLT 578

RESULT 12  
 ID W35754 standard; Protein: 133 AA.  
 AC W35754;  
 DT 18-FEB-1998 (first entry)  
 DE Maize rf2 gene partial protein sequence.  
 KW Maize; rf2 gene; male; fertility; cytoplasmic male sterility; CMS;  
 KW plant; hybrid; seed; pollination.  
 OS Zea mays.  
 FH Key Location/Qualifiers  
 FT Misc\_difference 32 /label= Unknown  
 FT /note= "Encoded by GNT"  
 FT Misc\_difference 56 /note= "Encoded by GGN"  
 FT Misc\_difference 118 /label= Unknown  
 FT /note= "Encoded by NTT"  
 FT Misc\_difference 121 /label= Unknown  
 FT /note= "Encoded by TGN"  
 FT Misc\_difference 123 /label= Unknown  
 FT /note= "Encoded by NCT"  
 FT Misc\_difference 126 /note= "Encoded by GTN"  
 FT Misc\_difference 131 /label= Unknown  
 FT /note= "Encoded by NAA"  
 FT US5684242-A.  
 PN 04-NOV-1997.  
 PD 29-NOV-1994; 346611.  
 PR 29-NOV-1994; US-346611.  
 PA (IOWA ) UNIV IOWA STATE RES FOUND INC.  
 PA (USDA ) US SEC OF AGRIC.  
 PI Schnable PS, Wise RP;  
 DR WPI: 97-549040/50.  
 DR N-PSDB; T94754.  
 PT Nucleic acid that reverses cytoplasmic male sterility in plants -  
 PT for producing hybrid seed and variant CMS factors  
 PS Example 4; Column 25-26; 17pp; English.  
 CC The present sequence represents a protein encoded by part of the maize  
 CC rf2 gene, which is capable of restoring male fertility to a plant with  
 CC cytoplasmic male sterility (CMS). The nucleic acid is used for producing  
 CC hybrid seed by inserting the nucleic acid into a plant cell,  
 CC establishing a first plant from the plant cell, and using the first  
 CC plant to pollinate a second plant having a CMS factor that is  
 CC compensated by the action of the male fertility restorer gene in the  
 CC first plant. It is also used for producing variant CMS factors by  
 CC inserting the nucleic acid into plant cells, culturing the cells to  
 CC generate a first plant, and selecting a second plant which has CMS but

CC which produces fertile seed when crossed with the first plant.  
SQ Sequence 133 AA;

Query Match 0.8%; Score 7; DB 26; Length 133;  
Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 ktvtlcl 100  
|||||  
QY 636 KVTLEL 642

## RESULT 13

ID R88107 standard; Protein; 139 AA.  
AC R88107;  
DE Murine anti-Protein C Mab HPC-4 VH gamma protein.  
KW Epitope; activation; heavy chain; protein C; vitamin K; plasma protein;  
KW zymogen; cleavage; mouse; humanised antibody; variable region;  
KW light chain; inhibition; anticoagulant; coagulation; tumour.  
OS Mus musculus.  
FH Key Location/Qualifiers  
FT peptide 1..19  
FT peptide /note= "signal peptide"  
FT peptide 20..139  
FT peptide /note= "mature peptide"  
PN W09534652-A1.  
PD 21-DEC-1995.  
PE 09-JUN-1995; U07372.  
PR 10-JUN-1994; US-259321.  
PA (OKLA-) OKLAHOMA MED RES FOUND.  
PI Esmon CT, Rezaie A;  
DR WPI; 96-049681/05.  
DR N-PSDB; T09299.  
PT Calcium-binding monoclonal antibody immunoreactive with Protein C -  
PT inhibits Protein C anticoagulant activation by  
PT thrombin-thrombomodulin, e.g. for treating tumours  
PS Claim 2; Page 29; 41pp; English.  
CC This is the amino acid sequence of the heavy chain variable region from  
CC the murine anti-protein C monoclonal antibody HPC-4 which recognises  
CC the activation peptide region (R88106) of the heavy chain of protein C,  
CC a vitamin K-dependent plasma protein zymogen. Protein C is converted to  
CC activated protein C (APC) by cleavage between the Arg-Leu amino acid  
CC contained within the activation peptide sequence. HPC-4 prevents protein  
CC C activation to APC by binding to this region. The DNA sequences encoding  
CC the variable regions of the heavy and light chains of the antibody  
CC (T09299-302) were used to construct humanised antibodies using the PCR  
CC primers T09303-9. The humanised antibodies are useful as inhibitors of  
CC coagulation and can be used for the treatment of tumours by inhibiting  
CC the anticoagulant activity of APC by preventing conversion of protein C  
CC to APC.  
SQ Sequence 139 AA;

Query Match 0.8%; Score 7; DB 17; Length 139;  
Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 34 sqtltlt 40  
|||||  
QY 572 SQTLLT 578

## RESULT 14

ID W76251 standard; Protein; 182 AA.  
AC W76251;  
DE Human TACE-like protein.  
KW TACE; TNF-alpha converting enzyme; tumour necrosis factor; screening;  
KW metalloproteinase; antagonist; diagnosis; overexpression; infection;  
KW inflammation; immune system; neurological disease.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..15

FT Protein /label= signal  
FT 16..182  
FT /label= TACE\_like\_protein  
FT 169..182  
FT /label= transmembrane\_region  
PN W09831818-A2.  
PD 23-JUL-1998.  
PE 20-JAN-1998; U00783.  
PR 01-AUG-1997; US-054541.  
PR 21-JAN-1997; US-034205.  
PR 13-JUN-1997; US-049607.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Brewer L, Gentz R, Ni J, Rosen CA, Ruben SM;  
DR WPI; 98-414114/35.  
DR N-PSDB; V61632.  
PT Isolated nucleic acid encoding human metallo-protease(s) - used for  
PT diagnosis, treatment and prevention of, e.g. cancer, inflammation,  
PT neurological disease and infections  
PS Claim 1a; Fig 1; 81pp; English.  
CC This sequence represents a novel human TACE-like protein which is a  
CC tumour necrosis factor (TNF) alpha converting enzyme and member of the  
CC metalloproteinase family. This protein can be used in assays to screen  
CC for agonists and antagonists and the nucleic acid is used as a probe  
CC for gene mapping, in situ hybridisation and detection of corresponding  
CC genes in human tissue, and as sources of probes and primers for  
CC diagnosis. The protein and its antigenic fragments are used to raise  
CC antibodies (Ab) (which can be used for diagnosis in usual immunoassays  
CC or for in vivo imaging) and to screen for (ant)agonists. Antagonists  
CC of this protein are used to treat disorders associated with  
CC overexpression of TNF-alpha, e.g. inflammation, immune system or  
CC neurological diseases, or infection.  
SQ Sequence 182 AA;

Query Match 0.8%; Score 7; DB 35; Length 182;  
Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 109 sicsglr 115  
|||||  
QY 384 SICSGLR 390

## RESULT 15

ID W35857 standard; protein; 194 AA.  
AC W35857;  
DE 27-APR-1998 (first entry)  
DE Human CD48 for use in T lymphocyte veto molecule.  
KW Human; CD48; T lymphocyte veto molecule; chimeric molecule;  
KW targeting polypeptide; suppression; immune response; treatment;  
KW autoimmune disease; allergy; immunological disorder;  
KW transplant rejection.  
OS Homo sapiens.  
PN W09737687-A1.  
PD 16-OCT-1997.  
PE 10-APR-1997; U05943.  
PR 10-APR-1996; US-630172.  
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
PI Staerz UD;  
DR WPI; 97-512419/47.  
PT T lymphocyte veto molecule comprising response cell activating  
PT protein - linked to molecule that targets stimulator cell marker,  
PT used for selective suppression of immune response, e.g. prevention  
PT of graft rejection or treatment of auto-immune disease  
PS Claim 37; Pages 72-73; 309pp; English.  
CC A novel T lymphocyte veto molecule is a chimeric molecule  
CC comprising a protein, e.g. the present sequence, linked to a  
CC targeting polypeptide that binds a molecule, which differentiates  
CC a host cell from a tissue graft cell, or selectively targets a  
CC stimulator cell involved in the autoimmune response.  
CC A veto molecule, in which the protein binds a molecule that targets  
CC stimulator cells, can be used to suppress an immune response and  
CC therefore treat autoimmune diseases, e.g. systemic lupus  
CC erythematosus, myasthenia gravis, rheumatoid arthritis, insulin

CC dependent diabetes mellitus, multiple sclerosis, coeliac disease,  
 CC autoimmune thyroiditis, Addison's or Grave's diseases and  
 CC rheumatoid arthritis, allergies and other immunological disorders.  
 CC Where the protein binds a molecule that differentiates graft and  
 CC host cells, the veto molecule can be used to reduce transplant  
 CC rejection. The veto molecule provides specific regulation of  
 CC particular stimulator cells that can kill graft cells or respond  
 CC to autoantigens, but leave other stimulator cells unaffected, e.g.  
 CC CD4 or CD8 positive cells can be regulated without one affecting  
 CC the other. The veto molecule can be administered locally to  
 CC minimise generalised immunosuppression.  
 SQ Sequence 194 AA:

Query Match 0.88; Score 7; DB 27; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 5.26e+01;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 64 qsgalyi 70

|||||

Qy 692 QSGALXI 698

Search completed: Sat Aug 28 14:36:45 1999  
 Job time : 100 secs.

**This Page Blank (uspio)**

```

MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm

Run on:      Sat Aug 28 14:33:29 1999;  MasPar time 35.62 Seconds
          1028.176 Million cell updates/sec
Tabular output not generated.

```

```

Title: >US-09-049-696-41
Description: (1-914) from US09049696.pep
Perfect Score: 914
Sequence: 1 MGPFSSVFILIHLEGAL.....GIHLKMKRWIGELOLSIA 914

```

Scoring table: TABLE unitprotatable

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

```
Database:
pir60
1:pir1 2:pir2 3:pir3 4:pir4
```

Statistics: Mean 3.987; Variance 0.423; scale 9.427

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query		ID	Description	Pred. No.
		Match	Length			
1	40	4.4	913	JG0168	gob-5 protein - Mouse	1.39e+00
2	11	1.2	794	T02171	hypothetical protein	3.18e+09
3	11	1.2	820	T02152	Lu-ECAM-1 protein - b	3.18e+09
4	11	1.2	905	T02205	Lu-ECAM-1 protein - b	3.18e+09
5	8	0.9	256	GFBP74	gene 14 protein - pha	6.56e+02
6	8	0.9	1320	S56279	proline dehydrogenase	6.56e+02
7	8	0.9	2573	D71614	hypothetical protein	6.56e+02
8	7	0.8	35	S61547	transferrin binding p	7.01e+00
9	7	0.8	48	A61331	trypsin (EC 3.4.21.4)	7.01e+00
10	7	0.8	76	S31653	ubiquitin - Giardia 1	7.01e+00
11	7	0.8	80	S75106	hypothetical protein	7.01e+00
12	7	0.8	121	G70074	hypothetical protein	7.01e+00
13	7	0.8	128	H70435	hypothetical protein	7.01e+00
14	7	0.8	142	G70646	hypothetical protein	7.01e+00
15	7	0.8	144	S74059	hypothetical protein	7.01e+00
16	7	0.8	145	S54237	Ig mu heavy chain V r	7.01e+00
17	7	0.8	153	S68406	vesicular integral me	7.01e+00
18	7	0.8	154	T70902	hypothetical protein	7.01e+00
19	7	0.8	173	H71679	hypothetical protein	7.01e+00
20	7	0.8	160	B53274	complement factor B s	7.01e+00
21	7	0.8	175	G69518	isochorismatase (enb	7.01e+00
22	7	0.8	175	S43894	pathogenesis-related	7.01e+00
23	7	0.8	179	ERAD34	early E3 20.1K glycop	7.01e+00

97	7	0.8	673	1	S01786	annexin VI - mouse	7.01e+00	170	6	0.7	131	2	H69397	iron-sulfur cluster b	3.77e+02
98	7	0.8	673	2	S52844	annexin VI - rat	7.01e+00	171	6	0.7	131	2	D24402	T-cell receptor alpha	3.77e+02
99	7	0.8	673	1	AIHUH68	annexin VI - human	7.01e+00	172	6	0.7	131	2	D24092	T-cell receptor alpha	3.77e+02
100	7	0.8	714	1	CIHUH	calpain (EC 3.4.22.17)	7.01e+00	173	6	0.7	131	2	A69071	conserved hypotheticala	3.77e+02
101	7	0.8	762	2	I59329	protein kinase (EC 2.7.0.1e+00)	7.01e+00	174	6	0.7	132	2	B70864	hypothetical protein	3.77e+02
102	7	0.8	762	1	A46590	protein kinase (EC 2.7.0.1e+00)	7.01e+00	175	6	0.7	133	2	A25777	T-cell receptor beta-	3.77e+02
103	7	0.8	762	2	S68217	protein kinase (EC 2.7.0.1e+00)	7.01e+00	176	6	0.7	134	2	H50908	hypothetical 15.1 kd	3.77e+02
104	7	0.8	763	1	A40831	gag-akt polyprotein -	7.01e+00	177	6	0.7	135	1	CBRT5M	cytochrome b5, outer	3.77e+02
105	7	0.8	823	2	H71300	probable DNA ligase (	7.01e+00	178	6	0.7	136	2	I51222	neurotrophin receptor	3.77e+02
106	7	0.8	866	2	D44234	fibrinogen alpha chain	7.01e+00	179	6	0.7	141	1	S38389	T-cell receptor beta	3.77e+02
107	7	0.8	900	1	GNNYMW	genome polyprotein -	7.01e+00	180	6	0.7	142	1	BVECTR	tolR protein - Escher	3.77e+02
108	7	0.8	932	2	JCS953	inter-alpha-inhibitor	7.01e+00	181	6	0.7	143	2	S6321	T-cell receptor delta	3.77e+02
109	7	0.8	996	2	S71750	import intermediate-a	7.01e+00	182	6	0.7	144	2	S6322	T-cell receptor delta	3.77e+02
110	7	0.8	999	1	IJHUG3	desmoglein 3 precursor	7.01e+00	183	6	0.7	145	2	S3554	probable membrane pro	3.77e+02
111	7	0.8	1136	2	C71110	hypothetical protein	7.01e+00	184	6	0.7	145	2	S72620	replication protein -	3.77e+02
112	7	0.8	2150	1	GNNYH2	genome polyprotein -	7.01e+00	185	6	0.7	146	1	HDWKDU	hemoglobin delta chain	3.77e+02
113	7	0.8	2157	2	S71461	proline-rich protein	7.01e+00	186	6	0.7	146	1	HDWKTB	hemoglobin delta chain	3.77e+02
114	7	0.8	2290	1	GNNYE	genome polyprotein -	7.01e+00	187	6	0.7	146	1	HDWKTB	hemoglobin delta chain	3.77e+02
115	7	0.8	2292	2	S55401	capsid polyprotein pr	7.01e+00	188	6	0.7	147	2	S6122	phosphotransferase sy	3.77e+02
116	7	0.8	2292	1	S35961	capsid polyprotein pr	7.01e+00	189	6	0.7	149	1	NEBP37	endodeoxyribonuclease	3.77e+02
117	7	0.8	2292	1	GNNYEB	genome polyprotein -	7.01e+00	190	6	0.7	150	2	S73725	probable lipoprotein	3.77e+02
118	7	0.8	2292	1	GNNYED	genome polyprotein -	7.01e+00	191	6	0.7	151	2	S65759	hypothetical protein	3.77e+02
119	7	0.8	4910	2	S64942	probable membrane pro	7.01e+00	192	6	0.7	151	2	S37812	hypothetical protein	3.77e+02
120	6	0.7	24	2	S55764	cathepsin G (EC 3.4.2	3.77e+02	193	6	0.7	151	2	S7322	probable heme-binding	3.77e+02
121	6	0.7	24	2	S16111	cathepsin G (EC 3.4.2	3.77e+02	194	6	0.7	151	2	S1987	probable heme-binding	3.77e+02
122	6	0.7	37	2	S03570	trypsin (EC 3.4.21.4)	3.77e+02	195	6	0.7	152	2	S73603	hypothetical protein	3.77e+02
123	6	0.7	40	2	S33409	Ig heavy chain V regi	3.77e+02	196	6	0.7	153	2	I73635	neuronal olfactomedin	3.77e+02
124	6	0.7	48	2	B21164	regulatory protein ro	3.77e+02	197	6	0.7	155	1	QOBCP1	ybeA protein - Escher	3.77e+02
125	6	0.7	50	2	P00026	photosystem I 9K prot	3.77e+02	198	6	0.7	157	2	H69553	conserved hypotheticala	3.77e+02
126	6	0.7	53	2	A46227	cement precursor prot	3.77e+02	199	6	0.7	157	2	A31105	finbrial protein prec	3.77e+02
127	6	0.7	60	2	S62445	hypothetical protein	3.77e+02	200	6	0.7	158	2	D70438	hypothetical protein	3.77e+02
128	6	0.7	62	2	A33717	nitrate reductase (ni	3.77e+02	201	6	0.7	159	2	S76716	hypothetical protein	3.77e+02
129	6	0.7	63	1	RGSCRE	regulatory protein ro	3.77e+02	202	6	0.7	160	2	H69804	transcription regulat	3.77e+02
130	6	0.7	68	1	PWU8	H+-transporting ATP s	3.77e+02	203	6	0.7	162	2	C34285	NADH dehydrogenase (u	3.77e+02
131	6	0.7	71	2	D27795	Na+/K+-exchanging ATP	3.77e+02	204	6	0.7	162	2	D69381	hypothetical protein	3.77e+02
132	6	0.7	72	2	G64130	hypothetical protein	3.77e+02	205	6	0.7	164	2	D69783	transcription regulat	3.77e+02
133	6	0.7	77	2	D70798	probable pPE protein	3.77e+02	206	6	0.7	164	2	S74345	hypothetical protein	3.77e+02
134	6	0.7	80	2	C63204	hypothetical protein	3.77e+02	207	6	0.7	165	2	P50261	calcineurin regulator	3.77e+02
135	6	0.7	85	2	F69933	hypothetical protein	3.77e+02	208	6	0.7	166	2	A4906	hypothetical protein	3.77e+02
136	6	0.7	91	2	A71968	hypothetical protein	3.77e+02	209	6	0.7	166	2	A42858	myosin light chain-2	3.77e+02
137	6	0.7	92	2	D78840	Ig kappa chain V regi	3.77e+02	210	6	0.7	168	2	S23348	H+-transporting ATP s	3.77e+02
138	6	0.7	92	2	S76807	hypothetical protein	3.77e+02	211	6	0.7	170	2	G71046	probable ferredoxin o	3.77e+02
139	6	0.7	95	2	D69269	ferredoxin (fdx-1) ho	3.77e+02	212	6	0.7	172	1	WMBEF5	UL45 protein - human	3.77e+02
140	6	0.7	95	2	S07110	T-cell receptor delta	3.77e+02	213	6	0.7	174	2	S06303	hypothetical protein	3.77e+02
141	6	0.7	96	2	S26924	Ig heavy chain V regi	3.77e+02	214	6	0.7	174	2	C69443	hypothetical protein	3.77e+02
142	6	0.7	97	2	G69122	hypothetical protein	3.77e+02	215	6	0.7	175	2	S06193	sflag protein precurs	3.77e+02
143	6	0.7	98	2	T01987	caffeoyl-CoA 3-O-meth	3.77e+02	216	6	0.7	175	2	C70453	translation initiatio	3.77e+02
144	6	0.7	99	2	S21350	hypothetical protein	3.77e+02	217	6	0.7	175	2	S56188	probable membrane pro	3.77e+02
145	6	0.7	104	2	F70014	conserved hypotheticala	3.77e+02	218	6	0.7	176	2	F1803	hypothetical protein	3.77e+02
146	6	0.7	106	1	A35135	thioredoxin - Rhodoba	3.77e+02	219	6	0.7	177	2	S34207	hypothetical protein	3.77e+02
147	6	0.7	107	2	B61396	collagen alpha 1(IX)	3.77e+02	220	6	0.7	177	2	E3852	ribosomal protein L5	3.77e+02
148	6	0.7	109	2	S07509	gene 4.2 protein - ph	3.77e+02	221	6	0.7	178	2	E69389	streptavidin v1 precu	3.77e+02
149	6	0.7	110	2	A55991	effector cell protein	3.77e+02	222	6	0.7	180	2	S14019	streptavidin precurs	3.77e+02
150	6	0.7	110	2	B24092	T-cell receptor alpha	3.77e+02	223	6	0.7	180	2	S1966	probable membrane pro	3.77e+02
151	6	0.7	111	2	S26255	T-cell receptor beta	3.77e+02	224	6	0.7	182	2	S4240	yop20 protein - yers	3.77e+02
152	6	0.7	111	2	S26256	T-cell receptor beta	3.77e+02	225	6	0.7	183	2	S57284	streptavidin v2 precu	3.77e+02
153	6	0.7	111	2	E69518	hypothetical protein	3.77e+02	226	6	0.7	183	2	S7285	streptavidin precurs	3.77e+02
154	6	0.7	113	2	S43109	acidic ribosomal prot	3.77e+02	227	6	0.7	183	2	A23513	calcium-binding prote	3.77e+02
155	6	0.7	113	2	S76792	hypothetical protein	3.77e+02	228	6	0.7	184	2	A24479	calcium-binding prote	3.77e+02
156	6	0.7	114	2	D70469	hypothetical protein	3.77e+02	229	6	0.7	184	2	A71230	hypothetical protein	3.77e+02
157	6	0.7	116	2	JCS188	sporulation-specific	3.77e+02	230	6	0.7	184	2	F59273	cytochrome-c oxidase	3.77e+02
158	6	0.7	119	2	C64174	hypothetical protein	3.77e+02	231	6	0.7	184	2	A37041	homeotic protein Ahox	3.77e+02
159	6	0.7	119	1	GIHWDW	Ig heavy chain V-II r	3.77e+02	232	6	0.7	185	2	T01953	calcium-binding prote	3.77e+02
160	6	0.7	120	2	JC2475	growth hormone regula	3.77e+02	233	6	0.7	190	2	T01953	hypothetical protein	3.77e+02
161	6	0.7	120	2	G70024	conserved hypotheticala	3.77e+02	234	6	0.7	192	2	J50373	hypothetical 22.5k pr	3.77e+02
162	6	0.7	121	2	S26290	trbA protein - Escher	3.77e+02	235	6	0.7	193	1	E38944	flagellin B3 precurs	3.77e+02
163	6	0.7	124	1	F6PMV4	vicillin, 14K componen	3.77e+02	236	6	0.7	194	1	REPCR5	resolvase - Escherich	3.77e+02
164	6	0.7	124	2	A53055	H+-transporting ATPas	3.77e+02	237	6	0.7	194	2	D69885	chorion protein s19 -	3.77e+02
165	6	0.7	124	2	JC4849	H+-transporting ATPas	3.77e+02	238	6	0.7	196	2	S06613	hibernation-related p	3.77e+02
166	6	0.7	124	2	A49002	Ig heavy chain V regi	3.77e+02	239	6	0.7	196	2	A48150	probable DNA-directed	3.77e+02
167	6	0.7	125	2	S04503	pancreatic ribonuclea	3.77e+02	240	6	0.7	198	2	S77811	anthranilate synthase	3.77e+02
168	6	0.7	126	2	S00790	hypothetical protein	3.77e+02	241	6	0.7	198	2	S35125	hypothetical protein	3.77e+02
169	6	0.7	129	2	A42692	T-cell receptor alpha	3.77e+02	242	6	0.7	200	2	S75591	hypothetical protein	3.77e+02



243	6	0.7	201	2	B71710	hypothetical protein	3.77e+02
244	6	0.7	201	2	S53511	peroxinectin-like pro	3.77e+02
245	6	0.7	203	2	A32009	sporulation repressor	3.77e+02
246	6	0.7	203	2	Jc4962	rab protein 30 - huma	3.77e+02
247	6	0.7	203	2	A36886	surface protein PAG n	3.77e+02
248	6	0.7	206	2	S58597	NADPH dehydrogenase (	3.77e+02
249	6	0.7	206	1	A55169	pre-r-cell receptor a	3.77e+02
250	6	0.7	206	2	S73867	hypothetical protein	3.77e+02
251	6	0.7	207	2	D70501	hypothetical protein	3.77e+02
252	6	0.7	207	2	A69446	conserved hypotheticala	3.77e+02
253	6	0.7	207	2	D64601	phosphoserine phosphat	3.77e+02
254	6	0.7	208	2	C70899	probable guanylate ki	3.77e+02
255	6	0.7	208	2	S28675	hypothetical protein	3.77e+02
256	6	0.7	208	2	S03615	glutathione transfera	3.77e+02
257	6	0.7	208	2	D45734	HisG - Lactococcus la	3.77e+02
258	6	0.7	210	2	C70528	hypothetical protein	3.77e+02
259	6	0.7	211	2	I53020	G-0/G-1 switch regula	3.77e+02
260	6	0.7	212	2	Jc4511	pyroglutamyl-peptida	3.77e+02
261	6	0.7	213	1	HRVOY	hemoglobin beta chain	3.77e+02
262	6	0.7	213	2	D70416	phosphoglycolate phos	3.77e+02
263	6	0.7	214	2	G70971	hypothetical protein	3.77e+02
264	6	0.7	215	2	C37273	membrane glycoprotein	3.77e+02
265	6	0.7	215	2	S25358	H+-transporting ATPas	3.77e+02
266	6	0.7	218	2	H71233	hypothetical protein	3.77e+02
267	6	0.7	228	2	D64672	ABC transporter, ATP-	3.77e+02
268	6	0.7	228	2	G71842	probable ABC transpor	3.77e+02
269	6	0.7	229	1	TRB07R	trypsin (EC 3.4.21.4)	3.77e+02
270	6	0.7	229	2	Jc0518	ribonuclease M5 (EC 3	3.77e+02
271	6	0.7	230	2	S33161	Ig kappa chain - shee	3.77e+02
272	6	0.7	231	2	G71201	probable triosephosph	3.77e+02
273	6	0.7	232	2	S56275	probable membrane pro	3.77e+02
274	6	0.7	233	2	S29577	Ig light chain - rain	3.77e+02
275	6	0.7	234	2	A40710	CD30 ligand - human	3.77e+02
276	6	0.7	234	2	H64384	pentose-5-phosphate-3	3.77e+02
277	6	0.7	236	2	H64743	copper homeostasis pr	3.77e+02
278	6	0.7	237	2	C70543	hypothetical protein	3.77e+02
279	6	0.7	237	2	G5084	hypothetical protein	3.77e+02
280	6	0.7	237	2	S36343	opacity protein opa57	3.77e+02
281	6	0.7	237	2	S55378	serine proteinase - A	3.77e+02
282	6	0.7	244	2	B69381	conserved hypotheticala	3.77e+02
283	6	0.7	244	2	S72219	chymotrypsin B - Atla	3.77e+02
284	6	0.7	244	2	JN0703	carboxyl reductase (N	3.77e+02
285	6	0.7	246	2	B48350	infected-cell protein	3.77e+02
286	6	0.7	246	1	TRRT1	trypsin (EC 3.4.21.4)	3.77e+02
287	6	0.7	246	1	TRRT2	trypsin (EC 3.4.21.4)	3.77e+02
288	6	0.7	247	1	A25852	trypsin (EC 3.4.21.4)	3.77e+02
289	6	0.7	247	2	A27547	trypsin (EC 3.4.21.4)	3.77e+02
290	6	0.7	248	2	A59109	conserved hypotheticala	3.77e+02
291	6	0.7	248	1	LNH0F6	pulmonary surfactant	3.77e+02
292	6	0.7	248	2	A43520	natural killer cell p	3.77e+02
293	6	0.7	249	2	S54437	hemin binding protein	3.77e+02
294	6	0.7	250	2	B71064	hypothetical protein	3.77e+02
295	6	0.7	250	2	G64709	hypothetical protein	3.77e+02
296	6	0.7	250	2	S55493	serine proteinase pre	3.77e+02
297	6	0.7	250	2	A70678	hypothetical protein	3.77e+02
298	6	0.7	251	2	H70488	conserved hypotheticala	3.77e+02
299	6	0.7	251	2	S24634	NADH dehydrogenase (u	3.77e+02
300	6	0.7	251	1	H5HBC	hemoglobin beta C(NA)	3.77e+02
301	6	0.7	252	2	S54838	cyclase - Rhodobacter	3.77e+02
302	6	0.7	253	2	Jc4594	cell surface RAE-1 pr	3.77e+02
303	6	0.7	253	2	C70317	HisF (cyclase) - Aqu	3.77e+02
304	6	0.7	253	2	A35654	hemoglobin linker cha	3.77e+02
305	6	0.7	256	2	S01045	hemolymph 30K protein	3.77e+02
306	6	0.7	256	2	A40169	protein-tyrosine-phos	3.77e+02
307	6	0.7	261	2	S75631	cyclase hisF - Synech	3.77e+02
308	6	0.7	261	2	S59631	glutathione transfera	3.77e+02
309	6	0.7	261	2	S16802	cyclase hisF - Azospi	3.77e+02
310	6	0.7	262	2	F71123	hypothetical protein	3.77e+02
311	6	0.7	262	2	S57276	14-3-3 protein homolo	3.77e+02
312	6	0.7	263	2	S47537	chymotrypsin (EC 3.4.	3.77e+02
313	6	0.7	263	2	JN0817	beta hemolysin precu	3.77e+02
314	6	0.7	263	2	B43727	cell division initiat	3.77e+02
315	6	0.7	264	2	S62909	ubiquitin precursor -	3.77e+02

2	2	S01922	hypothetical protein	3.77e+02
2	2	S75053	hypothetical protein	3.77e+02
2	2	F36914	molybdate-binding pro	3.77e+02
2	2	G64548	hypothetical protein	3.77e+02
2	2	A29119	chlorophyll a/b-bind	3.77e+02
2	2	B64644	hypothetical protein	3.77e+02
2	2	S06029	cytochrome-c oxidase	3.77e+02
2	2	D69127	hypothetical protein	3.77e+02
2	2	A36704	bifunctional folate b	3.77e+02
2	2	JE0274	connexin 31 - human	3.77e+02
1	1	LEWEC6	H+-transporting ATP s	3.77e+02
2	2	C69352	imidazoleglycerol-pho	3.77e+02
2	2	D42053	gap junction protein	3.77e+02
2	2	A70683	probable cy5w - Mycob	3.77e+02
2	2	E00572	hypothetical protein	3.77e+02
2	2	QOKWTA	hypothetical 32K prot	3.77e+02
2	2	D64201	methylentetrahydrofo	3.77e+02
2	2	S02506	nifu protein - Klebsi	3.77e+02
2	2	H70859	probable Enoyl-CoA hy	3.77e+02
2	2	S74967	UDP-3-O-acetyl N-acetyl	3.77e+02
2	2	A33604	rfs protein - Salmon	3.77e+02
2	2	T02471	hypothetical protein	3.77e+02
2	2	S76879	hypothetical protein	3.77e+02
2	2	A48488	saccharide biosynthes	3.77e+02
2	2	C70784	probable cyoA - Mycob	3.77e+02
2	2	S68216	phosphoprotein phosph	3.77e+02
2	2	S17820	protoporphyrin IX mag	3.77e+02
2	2	A41489	flagellar filament 31	3.77e+02
2	2	E70151	chemotaxis protein me	3.77e+02
2	2	A05235	hypothetical protein	3.77e+02
2	2	A55980	opsin-related retinal	3.77e+02
2	2	B42822	cyclin D3 - human	3.77e+02
2	2	B70319	conserved hypotheticala	3.77e+02
2	2	JC4012	cyclin D3 - rat	3.77e+02
2	2	S04619	beta-lactamase (EC 3.	3.77e+02
1	1	RDSGXX	ferredoxin--NADP+ red	3.77e+02
2	2	JE0174	ADP-ribosyl-nitrogena	3.77e+02
2	2	JT0536	frizzled protein-2 -	3.77e+02
2	2	S55980	opsin-related retinal	3.77e+02
2	2	S51278	alternative respirato	3.77e+02
2	2	S55085	hypothetical protein	3.77e+02
2	2	S39150	clathrin assembly pro	3.77e+02
2	2	B69840	hypothetical protein	3.77e+02
2	2	JQ2220	hydroxyproline-rich g	3.77e+02
2	2	F45045	phycobiliprotein homo	3.77e+02
2	2	T03599	phosphoprotein phosph	3.77e+02
2	2	F70356	transcription regulat	3.77e+02
2	2	T02851	hypothetical protein	3.77e+02
2	2	A44362	NADH dehydrogenase (u	3.77e+02
2	2	S32237	hypothetical protein	3.77e+02
2	2	S34687	hypothetical protein	3.77e+02
2	2	A32208	synaptophysin - bovin	3.77e+02
2	2	B69473	conserved hypotheticala	3.77e+02
2	2	JC6502	pectate lyase (EC 4.2	3.77e+02
2	2	S46239	ribosome-inactivating	3.77e+02
2	2	A42928	plasminogen activator	3.77e+02
2	2	A30916	coagulase/fibrinolysi	3.77e+02
2	2	S78254	DNA-directed RNA poly	3.77e+02
2	2	S60713	polygalacturonase-inh	3.77e+02
2	2	S26977	DNA-directed RNA poly	3.77e+02
2	2	JQ2381	S-locus-specific rece	3.77e+02
2	2	S77783	hypothetical protein	3.77e+02
2	2	G70313	ADP-heptose-LPS hepto	3.77e+02
2	2	E71240	hypothetical protein	3.77e+02
2	2	S35145	carboxyphosphoenolp	3.77e+02
2	2	E70117	membrane fusion prote	3.77e+02
2	2	B70194	conserved hypotheticala	3.77e+02
2	2	S69157	ferredoxin--NADP+ red	3.77e+02
1	1	LNH0ER	IgE Fc receptor II, 1	3.77e+02
2	2	C70700	hypothetical protein	3.77e+02
2	2	B64382	formylmethanofuran de	3.77e+02
2	2	W2WLE	E2 protein - human pa	3.77e+02
2	2	F71066	hypothetical protein	3.77e+02

389	6	0.7	327	2	G65126	probable general secr	3.77e+02	462	6	0.7	381	2	C41317	probable lipopolysacc	3.77e+02
390	6	0.7	329	2	S28305	hypothetical protein	3.77e+02	463	6	0.7	382	2	S09320	homeotic protein lin-	3.77e+02
391	6	0.7	331	2	G711498	hypothetical protein	3.77e+02	464	6	0.7	384	2	B69233	N-ethylamine chlor	3.77e+02
392	6	0.7	330	2	S01964	readthrough protein -	3.77e+02	465	6	0.7	385	2	A71146	hypothetical protein	3.77e+02
393	6	0.7	331	1	TVHUTN	transcription factor	3.77e+02	466	6	0.7	385	2	S61566	BBP1 protein - yeast	3.77e+02
394	6	0.7	332	2	S41417	glycoprotein VP7 prec	3.77e+02	467	6	0.7	386	1	S31140	alcohol dehydrogenase	3.77e+02
395	6	0.7	332	2	S41416	glycoprotein VP7 prec	3.77e+02	468	6	0.7	392	2	F69544	conserved hypothetical	3.77e+02
396	6	0.7	332	2	S711224	xyloglucan endo-1,4-b	3.77e+02	469	6	0.7	392	2	S37905	H+-transporting ATPas	3.77e+02
397	6	0.7	332	1	VGXRCN	glycoprotein VP7 prec	3.77e+02	470	6	0.7	394	2	S36512	E2 protein - human pa	3.77e+02
398	6	0.7	334	3	A54138	acidic repetitive pro	3.77e+02	471	6	0.7	395	2	I77371	CD44R5 - human	3.77e+02
399	6	0.7	335	2	G711221	hypothetical protein	3.77e+02	472	6	0.7	395	2	A47701	aspartic proteinase A	3.77e+02
400	6	0.7	335	2	B39999	hypothetical protein	3.77e+02	473	6	0.7	396	2	G48327	COI intron 6 protein	3.77e+02
401	6	0.7	336	1	PRASAK	glutamyl endopeptidas	3.77e+02	474	6	0.7	398	1	G3MSM	Ig gamma-3 chain C re	3.77e+02
402	6	0.7	337	2	A53041	effector cell protein	3.77e+02	475	6	0.7	399	2	B70936	probable serine-threo	3.77e+02
403	6	0.7	337	2	S62963	glyceraldehyde-3-phos	3.77e+02	476	6	0.7	400	2	S02507	nifs protein - Klebsi	3.77e+02
404	6	0.7	337	2	E69001	probable glyceraldehy	3.77e+02	477	6	0.7	402	2	S55980	probable membrane pro	3.77e+02
405	6	0.7	337	1	DEYDGC	glyceraldehyde-3-phos	3.77e+02	478	6	0.7	402	2	C71440	hypothetical protein	3.77e+02
406	6	0.7	339	2	A35967	homeotic protein roug	3.77e+02	479	6	0.7	403	2	A53662	homeotic protein HB9	3.77e+02
407	6	0.7	339	2	S70218	sipD protein - Salmon	3.77e+02	480	6	0.7	403	2	G64427	hypothetical protein	3.77e+02
408	6	0.7	340	1	Q0BSE6	UL20 protein precursor	3.77e+02	481	6	0.7	403	1	SUASO	oryzin (EC 3.4.21.63)	3.77e+02
409	6	0.7	340	2	H70790	hypothetical protein	3.77e+02	482	6	0.7	404	1	Q0BE13	BMRF1 protein - human	3.77e+02
410	6	0.7	341	2	I61725	natural killer associ	3.77e+02	483	6	0.7	404	1	S03849	ribonucleoprotein la	3.77e+02
411	6	0.7	341	2	G70723	hypothetical protein	3.77e+02	484	6	0.7	406	2	JC4996	biphenyl dioxygenase	3.77e+02
412	6	0.7	342	2	E69581	acetoacetyl protein	3.77e+02	485	6	0.7	407	2	I52703	42K membrane glycopro	3.77e+02
413	6	0.7	342	2	S04000	hypothetical protein	3.77e+02	486	6	0.7	407	2	A56607	arrestin homolog - ml	3.77e+02
414	6	0.7	344	2	T00110	integral membrane ace	3.77e+02	487	6	0.7	408	1	A31888	ribonucleoprotein La	3.77e+02
415	6	0.7	344	2	H64777	malate dehydrogenase	3.77e+02	488	6	0.7	409	2	T03718	suppressor 2 protein	3.77e+02
416	6	0.7	345	2	A53138	gas1 homolog - human	3.77e+02	489	6	0.7	409	2	S76126	hypothetical protein	3.77e+02
417	6	0.7	346	2	G70105	hypothetical protein	3.77e+02	490	6	0.7	410	2	I39766	glutamate N-acetyltra	3.77e+02
418	6	0.7	346	2	S72222	ferredoxin--NADP+ red	3.77e+02	491	6	0.7	410	1	JN0810	benzene 1,2-dioxygena	3.77e+02
419	6	0.7	347	2	F22845	hypothetical protein	3.77e+02	492	6	0.7	411	2	T01085	DNA binding protein b	3.77e+02
420	6	0.7	348	2	JN0639	recombinational protein	3.77e+02	493	6	0.7	414	2	S43340	glyceraldehyde-3-phos	3.77e+02
421	6	0.7	349	2	G70728	hypothetical protein	3.77e+02	494	6	0.7	415	1	JC1494	ribonucleoprotein La	3.77e+02
422	6	0.7	351	2	B65114	hypothetical protein	3.77e+02	495	6	0.7	416	2	A69023	conserved hypothetical	3.77e+02
423	6	0.7	352	2	F69901	fatty-acid desaturase	3.77e+02	496	6	0.7	418	2	E71716	proline/betaine trans	3.77e+02
424	6	0.7	353	2	S36249	lipB protein - Pseudo	3.77e+02	497	6	0.7	420	2	C69532	conserved hypothetical	3.77e+02
425	6	0.7	353	2	C69960	Xaa-Pro dipeptidase h	3.77e+02	498	6	0.7	422	2	A71542	probable glutamate am	3.77e+02
426	6	0.7	353	2	S37138	protoporphyrin IX mag	3.77e+02	499	6	0.7	423	2	F69031	homoserine dehydrogen	3.77e+02
427	6	0.7	354	2	S52759	prephenate dehydrogen	3.77e+02	500	6	0.7	424	2	D71192	hypothetical protein	3.77e+02
428	6	0.7	356	2	G70398	hypothetical protein	3.77e+02	501	6	0.7	424	2	A71846	glycinamide ribonucle	3.77e+02
429	6	0.7	356	2	S32147	naringenin 3-dioxygen	3.77e+02	502	6	0.7	424	2	B64672	glycinamide ribonucle	3.77e+02
430	6	0.7	357	2	S21758	glutamic acid-specifi	3.77e+02	503	6	0.7	425	2	A38153	paired box protein Po	3.77e+02
431	6	0.7	358	2	I50506	gene wnt8b protein -	3.77e+02	504	6	0.7	427	2	G71272	probable nitrogen fix	3.77e+02
432	6	0.7	360	3	T01998	hypothetical protein	3.77e+02	505	6	0.7	427	2	A55496	apexin precursor - gu	3.77e+02
433	6	0.7	362	2	S45734	Hisc - Lactococcus la	3.77e+02	506	6	0.7	427	2	JQ2203	UL48h protein - Marek	3.77e+02
434	6	0.7	361	2	JH0417	cell adhesion molecucl	3.77e+02	507	6	0.7	427	2	S60742	alpha-transinducing f	3.77e+02
435	6	0.7	361	2	G02251	cell surface glycopro	3.77e+02	508	6	0.7	428	1	EHUO	Ig epsilon chain C re	3.77e+02
436	6	0.7	361	2	A32376	lymphocyte surface an	3.77e+02	509	6	0.7	430	2	A56002	pre-B-cell leukemia t	3.77e+02
437	6	0.7	361	2	F64355	conserved hypothetical	3.77e+02	510	6	0.7	431	2	D47553	phor protein - Klebsi	3.77e+02
438	6	0.7	362	2	S45987	ribosomal protein L4 .	3.77e+02	511	6	0.7	433	2	JU0396	fusaric acid resistan	3.77e+02
439	6	0.7	363	1	CBUTB	ubiquinol--cytochrome	3.77e+02	512	6	0.7	434	2	C69397	hypothetical protein	3.77e+02
440	6	0.7	363	2	S69444	probable serine/threo	3.77e+02	513	6	0.7	434	2	JC1039	phosphopyruvate hydra	3.77e+02
441	6	0.7	365	2	A39481	serum response factor	3.77e+02	514	6	0.7	434	1	NOHUG	phosphopyruvate hydra	3.77e+02
442	6	0.7	366	2	F71269	probable tRNA pseudou	3.77e+02	515	6	0.7	436	2	E69371	bile acid-inducible o	3.77e+02
443	6	0.7	366	2	G71179	probable acetylornith	3.77e+02	516	6	0.7	436	2	A70409	hypothetical protein	3.77e+02
444	6	0.7	367	2	S53898	probable membrane pro	3.77e+02	517	6	0.7	437	2	I57942	5-hydroxytryptamine r	3.77e+02
445	6	0.7	367	2	I39172	cyclin-dependent kina	3.77e+02	518	6	0.7	438	2	A47702	glucan 1,3-beta-gluc	3.77e+02
446	6	0.7	368	2	B45610	aldolase ALDO-1 - pla	3.77e+02	519	6	0.7	438	2	JC5548	cell division protein	3.77e+02
447	6	0.7	368	2	A45610	fructose-bisphosphate	3.77e+02	520	6	0.7	439	2	S20489	S-locus-specific glyc	3.77e+02
448	6	0.7	369	2	S74017	hypothetical protein	3.77e+02	521	6	0.7	440	2	B71153	hypothetical protein	3.77e+02
449	6	0.7	370	2	A25004	keratin, 53K type II	3.77e+02	522	6	0.7	441	2	F70759	hypothetical protein	3.77e+02
450	6	0.7	370	2	G70464	GRP-binding protein H	3.77e+02	523	6	0.7	441	2	B69004	hypothetical protein	3.77e+02
451	6	0.7	370	2	B35255	chloromuconate cycloi	3.77e+02	524	6	0.7	442	2	T03172	helicase homolog 086L	3.77e+02
452	6	0.7	371	2	F69429	asparagine synthetase	3.77e+02	525	6	0.7	442	2	S76073	hypothetical protein	3.77e+02
453	6	0.7	372	1	VVPK1	coat protein VP1 - mo	3.77e+02	526	6	0.7	443	2	D71405	probable protein kina	3.77e+02
454	6	0.7	374	2	S69339	Ig heavy chain V regi	3.77e+02	527	6	0.7	444	2	T01721	hypothetical protein	3.77e+02
455	6	0.7	374	2	S72664	Ig heavy chain V regi	3.77e+02	528	6	0.7	445	2	S26601	transcription activat	3.77e+02
456	6	0.7	374	2	S57750	naringenin 3-dioxygen	3.77e+02	529	6	0.7	446	2	D71418	hypothetical proteint	3.77e+02
457	6	0.7	374	2	T02315	hypothetical protein	3.77e+02	530	6	0.7	447	2	F64706	UDP-MurNac-tripeptide	3.77e+02
458	6	0.7	375	2	F69551	coenzyme POQ synthesi	3.77e+02	531	6	0.7	447	2	D71812	udp-n-acetylmuramyl-t	3.77e+02
459	6	0.7	377	1	PEMOCJ	gas2ricin (EC 3.4.23	3.77e+02	532	6	0.7	448	1	VH1HN1	nucleocapsid protein	3.77e+02
460	6	0.7	380	2	T02342	DNA-binding protein -	3.77e+02	533	6	0.7	448	2	C64437	probable urease - Met	3.77e+02
461	6	0.7	380	2	S55923	cysteine proteinase (	3.77e+02	534	6	0.7	448	2	JQ1173	nucleocapsid protein	3.77e+02

535	6	0.7	449	2	I39489	L-sorbose dehydroge	3.77e+02	608	518	2	G64986	hypothetical protein	3.77e+02
536	6	0.7	450	2	S38114	hypothetical protein	3.77e+02	609	519	2	S31136	nitrogen fixation reg	3.77e+02
537	6	0.7	451	2	WABT5	gene D10 protein - ph	3.77e+02	610	520	1	S02505	nitrogenase (EC 1.18.	3.77e+02
538	6	0.7	452	2	S77155	UDP-N-acetylmuramoyl	3.77e+02	611	520	2	S35575	myosin heavy chain, c	3.77e+02
539	6	0.7	453	2	C31933	fg mu chain C region	3.77e+02	612	521	2	S67491	phosphate transport p	3.77e+02
540	6	0.7	454	2	I61727	natural killer associ	3.77e+02	613	521	2	B70182	lysine--trNA ligase (	3.77e+02
541	6	0.7	455	2	G01923	KIR (cl-5) NK recept	3.77e+02	614	523	2	S57811	leucyl aminopeptidase	3.77e+02
542	6	0.7	456	2	A49493	tyrosine phenol-lyase	3.77e+02	615	524	2	A44982	flagellin - Shigella	3.77e+02
543	6	0.7	456	2	S26357	tyrosine phenol-lyase	3.77e+02	616	524	2	E71525	probable ABC transport	3.77e+02
544	6	0.7	456	2	I52555	tyrosine phenol-lyase	3.77e+02	617	529	2	A41137	heat shock transcript	3.77e+02
545	6	0.7	456	2	S47924	MGAL protein - yeast	3.77e+02	618	530	2	A45334	3',5'-cyclic-nucleoti	3.77e+02
546	6	0.7	458	1	S77662	thioredoxin reductase	3.77e+02	619	533	2	E69154	CTP synthase - Methan	3.77e+02
547	6	0.7	458	2	G69123	conserved hypothetical	3.77e+02	620	535	2	C46243	epidermal growth fact	3.77e+02
548	6	0.7	459	2	S68519	tub protein, testis -	3.77e+02	621	535	2	S78598	D-ribulokinase (EC 2.	3.77e+02
549	6	0.7	460	2	I57546	Rabin3 - rat	3.77e+02	622	538	2	H69205	chaperonin - Methanob	3.77e+02
550	6	0.7	461	2	B69576	alkaline phosphatase	3.77e+02	623	543	2	A53790	cytochrome P450-EF -	3.77e+02
551	6	0.7	461	2	S77728	probable transport pr	3.77e+02	624	544	2	S24915	gene 18 protein - pha	3.77e+02
552	6	0.7	461	1	KXHU	protein C (activated)	3.77e+02	625	544	2	S65231	tRNA-pseudouridine sy	3.77e+02
553	6	0.7	461	2	C64733	protein transport pro	3.77e+02	626	545	2	F64579	adenine specific DNA	3.77e+02
554	6	0.7	462	2	T02726	probable protein kina	3.77e+02	627	545	2	F71932	type II DNA modifcat	3.77e+02
555	6	0.7	462	1	WTHUB	semenogelin I precurs	3.77e+02	628	546	2	T02029	DNA-binding protein p	3.77e+02
556	6	0.7	462	2	A64416	adenylosuccinate lyas	3.77e+02	629	548	2	E70546	hypothetical protein	3.77e+02
557	6	0.7	464	2	JE0092	mitochondrial complex	3.77e+02	630	552	2	H69126	chaperonin - Methanob	3.77e+02
558	6	0.7	464	2	A39362	NADH dehydrogenase (u	3.77e+02	631	553	1	DJBPD1	DNA-directed DNA poly	3.77e+02
559	6	0.7	464	2	G71825	hypothetical protein	3.77e+02	632	554	2	S34607	carboxylesterase (EC	3.77e+02
560	6	0.7	465	2	H69120	hypothetical protein	3.77e+02	633	555	2	C70177	beta-glucosidase homo	3.77e+02
561	6	0.7	466	2	C64690	conserved hypothetical	3.77e+02	634	555	2	S27163	alpha,alpha-trehalase	3.77e+02
562	6	0.7	466	2	I39707	hydrolase - Agrobacte	3.77e+02	635	555	1	SYHQA	malate synthase (EC 4	3.77e+02
563	6	0.7	467	2	C63150	Gram-negative bacteri	3.77e+02	636	557	2	S62002	hypothetical protein	3.77e+02
564	6	0.7	469	2	A48788	leucyl aminopeptidase	3.77e+02	637	558	2	S32914	pilF protein - Neisse	3.77e+02
565	6	0.7	470	2	I38026	MLN 62 protein - huma	3.77e+02	638	560	2	A40044	PML-1 protein - human	3.77e+02
566	6	0.7	471	2	S78597	argininosuccinate lya	3.77e+02	639	561	2	S63388	probable membrane pro	3.77e+02
567	6	0.7	472	2	D71076	probable glutamate sy	3.77e+02	640	566	2	E64774	probable membrane pro	3.77e+02
568	6	0.7	473	2	S51356	probable membrane pro	3.77e+02	641	566	1	B41648	protein-tyrosine-phos	3.77e+02
569	6	0.7	475	2	A69149	O-antigen transporter	3.77e+02	642	567	2	A31167	RNA-directed RNA poly	3.77e+02
570	6	0.7	476	2	A56510	dolichyl-diphosphooli	3.77e+02	643	567	2	D70769	hypothetical protein	3.77e+02
571	6	0.7	476	2	A39406	dihydrolipoamide dehy	3.77e+02	644	567	2	A40899	gag polyprotein - Chi	3.77e+02
572	6	0.7	477	1	DEAVHL	dihydrolipoamide dehy	3.77e+02	645	568	2	S57830	glucose-6-phosphate 1	3.77e+02
573	6	0.7	478	2	S62554	glutathione synthase	3.77e+02	646	568	1	A05225	gamma-glutamyltransfe	3.77e+02
574	6	0.7	479	2	A13303	conserved hypothetical	3.77e+02	647	569	2	S57812	leucyl aminopeptidase	3.77e+02
575	6	0.7	482	2	G64771	yajK protein - Escher	3.77e+02	648	569	2	T00445	vacuolar protein-sort	3.77e+02
576	6	0.7	483	2	B71537	probable N-acetylmura	3.77e+02	649	569	2	A41806	glucose-6-phosphate 1	3.77e+02
577	6	0.7	485	2	I73637	neuronal olfactomedin	3.77e+02	650	569	2	S64957	asparagillopepsin I (E	3.77e+02
578	6	0.7	486	1	FOLJEV	gag polyprotein - equ	3.77e+02	651	570	2	S35069	transcription factor	3.77e+02
579	6	0.7	490	2	H69858	cation ABC transporte	3.77e+02	652	574	2	B35149	lpah protein - Shigel	3.77e+02
580	6	0.7	491	2	A40383	3',5'-cyclic-nucleoti	3.77e+02	653	574	2	I66868	cdc35B - rat	3.77e+02
581	6	0.7	491	2	B69104	conserved hypothetical	3.77e+02	654	574	2	E64414	hypothetical protein	3.77e+02
582	6	0.7	492	2	G64550	ATP-dependent RNA hel	3.77e+02	655	575	2	S35786	glycoprotein gE - bov	3.77e+02
583	6	0.7	493	2	S13330	CD44E protein, epithe	3.77e+02	656	576	2	A42336	protein-tyrosine-phos	3.77e+02
584	6	0.7	494	2	S21779	pacA protein - phage	3.77e+02	657	576	2	S69214	deformed epidermal au	3.77e+02
585	6	0.7	494	2	A27442	triacylglycerol lipas	3.77e+02	658	579	4	D40201	artfact-warning sequ	3.77e+02
586	6	0.7	497	2	JN0420	DNA-directed RNA poly	3.77e+02	659	583	2	S01496	lamin B - African cla	3.77e+02
587	6	0.7	498	1	VGBEGX	secreted glycoprotein	3.77e+02	660	583	2	T02045	kinase associated pro	3.77e+02
588	6	0.7	498	1	W2WL8	E2 protein - human pa	3.77e+02	661	585	2	S48951	hypothetical protein	3.77e+02
589	6	0.7	498	2	S77411	integral membrane pro	3.77e+02	662	586	2	JQ0184	old protein - phage P	3.77e+02
590	6	0.7	499	2	S47713	pita protein - Escher	3.77e+02	663	586	2	A45144	H+-transporting ATP s	3.77e+02
591	6	0.7	500	2	A36388	RNA-directed RNA poly	3.77e+02	664	589	2	A60198	Myl protein - human	3.77e+02
592	6	0.7	500	2	S56489	hypothetical 54.3k pr	3.77e+02	665	589	2	S42517	PML protein, splice f	3.77e+02
593	6	0.7	503	2	JC5321	cytochrome P450 monoo	3.77e+02	666	589	2	S44381	PML protein, splice f	3.77e+02
594	6	0.7	503	2	S62018	probable membrane pro	3.77e+02	667	589	2	F69426	chloride channel, pro	3.77e+02
595	6	0.7	503	2	S09640	variant surface glyco	3.77e+02	668	590	2	A40437	glutamic acid-rich pr	3.77e+02
596	6	0.7	505	2	S68518	tub protein, brain -	3.77e+02	669	591	2	A40684	early B-cell factor -	3.77e+02
597	6	0.7	507	2	E71551	probable monooxygenas	3.77e+02	670	591	1	NUZQF	glucose-6-phosphate 1	3.77e+02
598	6	0.7	508	2	A33378	fasciclin III precurs	3.77e+02	671	591	2	S77707	phosphoribosylaminoi	3.77e+02
599	6	0.7	510	2	A42750	insulinoma-associated	3.77e+02	672	592	2	A28652	H+-transporting ATP s	3.77e+02
600	6	0.7	510	2	S38657	sucrose transport pro	3.77e+02	673	593	2	B40045	probable transcriptio	3.77e+02
601	6	0.7	510	1	DERN2	NADH dehydrogenase (u	3.77e+02	674	594	2	S50611	hypothetical protein	3.77e+02
602	6	0.7	510	2	A64706	lipase-like protein -	3.77e+02	675	595	2	A38628	threonine dehydratase	3.77e+02
603	6	0.7	510	2	S65075	NADH dehydrogenase (u	3.77e+02	676	595	2	A46373	serine/threonine kina	3.77e+02
604	6	0.7	510	1	DEECCA	NAD(P)+ transhydrogen	3.77e+02	677	596	2	JN0895	succinate dehydrogena	3.77e+02
605	6	0.7	510	2	A71816	hypothetical protein	3.77e+02	678	597	4	E40201	artfact-warning sequ	3.77e+02
606	6	0.7	510	2	S68223	glutathione synthase	3.77e+02	679	597	2	A46050	thyroid/steroid recep	3.77e+02
607	6	0.7	515	2	T03717	GTP-binding protein S	3.77e+02	680	598	2	I64042	GTP-binding membrane	3.77e+02

681	6	0.7	598	2	A46225	Nur-related transcript	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
682	6	0.7	598	2	I84692	nuclear orphan receptor	3.77e+02	3.77e+02	peptidyl-dipeptidase	3.77e+02
683	6	0.7	599	2	Q2994	CDPK-related protein	3.77e+02	3.77e+02	photosystem I protein	3.77e+02
684	6	0.7	599	1	BVCELA	GTP-binding membrane	3.77e+02	3.77e+02	glycine--tRNA ligase	3.77e+02
685	6	0.7	602	2	JH0166	potassium voltage-gat	3.77e+02	3.77e+02	transcription factor	3.77e+02
686	6	0.7	605	2	JH0134	probable GTP-binding	3.77e+02	3.77e+02	lymphocyte homing rec	3.77e+02
687	6	0.7	607	1	QXMS5M	NADH dehydrogenase (u	3.77e+02	3.77e+02	82.3K alpha trans-ind	3.77e+02
688	6	0.7	607	1	NUUTB	glucose-6-phosphate i	3.77e+02	3.77e+02	ferriochrome-iron rece	3.77e+02
689	6	0.7	607	1	ABXL72	74K albumin precursor	3.77e+02	3.77e+02	prohormone-processing	3.77e+02
690	6	0.7	608	1	ABXL68	68K serum albumin pre	3.77e+02	3.77e+02	ATP-dependent clip pro	3.77e+02
691	6	0.7	609	2	S43009	hemagglutinin - rinde	3.77e+02	3.77e+02	probable lipoprotein	3.77e+02
692	6	0.7	609	1	HMZKA	hemagglutinin - rinde	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
693	6	0.7	612	1	WDSP15	gene D13 protein - ph	3.77e+02	3.77e+02	inorganic pyrophospha	3.77e+02
694	6	0.7	613	4	C42021	artifact-warring sequ	3.77e+02	3.77e+02	isoamylase (EC 3.2.1.1	3.77e+02
695	6	0.7	614	2	S11320	probable adenylate cy	3.77e+02	3.77e+02	excision repair prote	3.77e+02
696	6	0.7	616	2	S67456	ORC3 protein - yeast	3.77e+02	3.77e+02	androgen-regulated ep	3.77e+02
697	6	0.7	616	2	A40595	methylmalonyl-CoA mut	3.77e+02	3.77e+02	DNA mismatch repair p	3.77e+02
698	6	0.7	616	2	S64624	alpha-glucoside trans	3.77e+02	3.77e+02	helicase (EC 3.6.1.~)	3.77e+02
699	6	0.7	618	2	A55188	nucleolar protein NOP	3.77e+02	3.77e+02	DNA mismatch repair p	3.77e+02
700	6	0.7	620	2	S55086	probable membrane pro	3.77e+02	3.77e+02	infected cell protein	3.77e+02
701	6	0.7	621	2	S11510	neurofibromatosis-rel	3.77e+02	3.77e+02	cell division protein	3.77e+02
702	6	0.7	621	2	T01935	naringenin 3-dioxygen	3.77e+02	3.77e+02	probable iron-regulat	3.77e+02
703	6	0.7	622	2	S73452	hypothetical protein	3.77e+02	3.77e+02	iron-regulated outer	3.77e+02
704	6	0.7	627	4	A40201	artifact-warring sequ	3.77e+02	3.77e+02	X-Pro dipeptidyl-pept	3.77e+02
705	6	0.7	629	2	S31174	hypothetical protein	3.77e+02	3.77e+02	PML protein, splice f	3.77e+02
706	6	0.7	630	2	S37884	finger protein MSN4 -	3.77e+02	3.77e+02	PML protein, splice f	3.77e+02
707	6	0.7	631	2	E71933	hypothetical protein	3.77e+02	3.77e+02	bcsB protein - Acetob	3.77e+02
708	6	0.7	631	2	A64578	hypothetical protein	3.77e+02	3.77e+02	protein-tyrosine kina	3.77e+02
709	6	0.7	632	2	S58152	hypothetical protein	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
710	6	0.7	633	2	S19244	gene Myl protein - hu	3.77e+02	3.77e+02	DNA mismatch repair p	3.77e+02
711	6	0.7	633	1	ZPSCP2	penicillin-binding pr	3.77e+02	3.77e+02	probable membrane pro	3.77e+02
712	6	0.7	634	2	S28724	conserved hypothetical	3.77e+02	3.77e+02	CDC27 protein - human	3.77e+02
713	6	0.7	641	2	I50642	gammaAPP-B - chicken	3.77e+02	3.77e+02	H-NUC - human	3.77e+02
714	6	0.7	641	2	A40045	probable transcriptio	3.77e+02	3.77e+02	transcription factor	3.77e+02
715	6	0.7	642	2	S53434	plasma protein S prec	3.77e+02	3.77e+02	neurotrophin-3 recept	3.77e+02
716	6	0.7	643	2	A43423	dynein 74K chain, cyt	3.77e+02	3.77e+02	neurotrophin-3 recept	3.77e+02
717	6	0.7	643	2	S63681	hypothetical protein	3.77e+02	3.77e+02	probable membrane pro	3.77e+02
718	6	0.7	646	1	W2BEC8	68.6K capsid protein	3.77e+02	3.77e+02	S-receptor kinase (EC	3.77e+02
719	6	0.7	649	2	S67787	hypothetical protein	3.77e+02	3.77e+02	invasin - Yersinia en	3.77e+02
720	6	0.7	654	2	I38605	oviductal glycoprotei	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
721	6	0.7	657	2	S05517	lamin - chicken	3.77e+02	3.77e+02	DNA helicase/primase	3.77e+02
722	6	0.7	657	2	S25184	cspl protein - Coryne	3.77e+02	3.77e+02	DNA helicase/primase	3.77e+02
723	6	0.7	660	2	A64739	ferriochrome-iron tran	3.77e+02	3.77e+02	neurotrophin-3 recept	3.77e+02
724	6	0.7	660	2	B34469	pullulanase secretion	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
725	6	0.7	662	1	S12989	histidine decarboxyla	3.77e+02	3.77e+02	S-receptor kinase (EC	3.77e+02
726	6	0.7	664	2	I56171	B-cell adhesion prote	3.77e+02	3.77e+02	epidermal growth fact	3.77e+02
727	6	0.7	665	2	I50641	gammaAPP-A - chicken	3.77e+02	3.77e+02	B-cell adhesion prote	3.77e+02
728	6	0.7	671	2	A40692	signal recognition pa	3.77e+02	3.77e+02	KI domain interacting	3.77e+02
729	6	0.7	674	2	JC5104	transcription initiat	3.77e+02	3.77e+02	pre-mRNA processing p	3.77e+02
730	6	0.7	676	2	I50643	gammaAPP-C - chicken	3.77e+02	3.77e+02	tyrosine kinase C rec	3.77e+02
731	6	0.7	677	1	KXHUS	collagen alpha 2(IX)	3.77e+02	3.77e+02	1,4-alpha-glucan bran	3.77e+02
732	6	0.7	677	2	S23296	levanase (EC 3.2.1.65	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
733	6	0.7	677	2	A42886	hypothetical protein	3.77e+02	3.77e+02	calcium-binding protei	3.77e+02
734	6	0.7	679	2	S38011	probable transport pr	3.77e+02	3.77e+02	nitrate reductase (NA	3.77e+02
735	6	0.7	680	2	T03106	collagen alpha 1(X) c	3.77e+02	3.77e+02	lipoxigenase (EC 1.13	3.77e+02
736	6	0.7	680	1	CGHU1D	collagen alpha 1(X) c	3.77e+02	3.77e+02	replicative DNA helic	3.77e+02
737	6	0.7	684	2	A56154	Abl substrate ena (en	3.77e+02	3.77e+02	calphotin - fruit fly	3.77e+02
738	6	0.7	686	2	E71895	probable heavy-metal	3.77e+02	3.77e+02	glycoprotein B - equi	3.77e+02
739	6	0.7	688	2	S64671	polynucleotide phosph	3.77e+02	3.77e+02	glycoprotein B - equi	3.77e+02
740	6	0.7	689	2	S70661	transferrin-binding p	3.77e+02	3.77e+02	iota toxin component	3.77e+02
741	6	0.7	699	2	I37369	epican - human	3.77e+02	3.77e+02	glycoprotein B precu	3.77e+02
742	6	0.7	701	2	S73722	protein secretion fac	3.77e+02	3.77e+02	glycoprotein B precu	3.77e+02
743	6	0.7	701	2	F70155	Na+/H+ antiporter (na	3.77e+02	3.77e+02	UDPglucose 6-dehydrog	3.77e+02
744	6	0.7	701	2	S62460	hypothetical protein	3.77e+02	3.77e+02	clathrin assembly pro	3.77e+02
745	6	0.7	702	2	C64835	hypothetical protein	3.77e+02	3.77e+02	hypothetical protein	3.77e+02
746	6	0.7	705	2	S55133	heat shock protein HS	3.77e+02	3.77e+02	phosphoprotein, synap	3.77e+02
747	6	0.7	712	2	A48156	translation regulator	3.77e+02	3.77e+02	probable cation trans	3.77e+02
748	6	0.7	714	1	JG9408	phosphoribosylformylg	3.77e+02	3.77e+02	orphan G protein-coup	3.77e+02
749	6	0.7	715	2	GC4908	alkaline serine prote	3.77e+02	3.77e+02	inter-alpha-trypsin i	3.77e+02
750	6	0.7	716	2	A56628	phenylalanine ammonia	3.77e+02	3.77e+02		
751	6	0.7	719	2	T00266	hypothetical protein	3.77e+02	3.77e+02		
752	6	0.7	722	2	B71376	hypothetical protein	3.77e+02	3.77e+02		
753	6	0.7	727	2	G01792	CTCF - human	3.77e+02	3.77e+02		

827	6	0.7	914	2	S18942	hypothetical protein	3.77e+02	900	2	S22624	aggregation protein a	3.77e+02
828	6	0.7	915	2	S36337	clathrin assembly pro	3.77e+02	901	2	S41552	probable transcrip	3.77e+02
829	6	0.7	920	2	S52301	nitrate reductase (NA	3.77e+02	902	2	S55669	tegument protein 75-	3.77e+02
830	6	0.7	921	2	S42617	collagen alpha 1(IX)	3.77e+02	903	2	A36080	insulin receptor prec	3.77e+02
831	6	0.7	922	2	D70066	SNF2 helixase homolog	3.77e+02	904	6	T00257	hypothetical protein	3.77e+02
832	6	0.7	930	2	JX0368	inter-alpha-trypsin i	3.77e+02	905	6	I58157	peritaxin - rat	3.77e+02
833	6	0.7	932	2	S47597	mutL protein homolog	3.77e+02	906	6	A43336	microtubule-vesicle 1	3.77e+02
834	6	0.7	937	1	D3VZ2W	DNA-directed DNA poly	3.77e+02	907	6	S54471	probable membrane pro	3.77e+02
835	6	0.7	952	1	HXAD5	hexon protein - human	3.77e+02	908	6	T00333	hypothetical protein	3.77e+02
836	6	0.7	955	4	C40045	probable transcrip	3.77e+02	909	6	S22595	restin - human	3.77e+02
837	6	0.7	957	2	S36834	glycine dehydrogenase	3.77e+02	910	6	A45866	dehydroascorbate (EC 2.	3.77e+02
838	6	0.7	959	2	H69344	hypothetical protein	3.77e+02	911	6	A30588	140K adhesin precursor	3.77e+02
839	6	0.7	970	2	A41944	mitotic control prote	3.77e+02	912	6	A44027	165K myofibrillar M-b	3.77e+02
840	6	0.7	975	2	A42466	alpha-amylase (EC 3.2	3.77e+02	913	6	J01719	E2 glycoprotein precu	3.77e+02
841	6	0.7	985	2	JC4217	alpha-glucosidase (EC	3.77e+02	914	6	S58350	DNA-directed DNA poly	3.77e+02
842	6	0.7	987	1	J00016	6-phosphofructokinase	3.77e+02	915	6	A60026	cell communication-me	3.77e+02
843	6	0.7	991	2	I40229	arginyl endopeptidase	3.77e+02	916	6	B43855	high-molecular-weight	3.77e+02
844	6	0.7	992	2	S32902	6-phosphofructokinase	3.77e+02	917	6	A47341	amylopullulanase prec	3.77e+02
845	6	0.7	994	1	SVNCLM	leucine--trNA ligase	3.77e+02	918	6	S42839	tl6G12.5 protein - Ca	3.77e+02
846	6	0.7	1004	2	A39611	probable GTP-binding	3.77e+02	919	6	T02850	probable membrane pro	3.77e+02
847	6	0.7	1018	2	S73720	cytadherence accessor	3.77e+02	920	6	T01098	chloroplast outer env	3.77e+02
848	6	0.7	1021	2	I39207	leukocyte surface pro	3.77e+02	921	6	S45768	mitotic spindle prote	3.77e+02
849	6	0.7	1022	2	I53078	homeotic gene regulat	3.77e+02	922	6	A44811	glucosyltransferase (	3.77e+02
850	6	0.7	1023	2	JC4013	major acidic nuclear	3.77e+02	923	6	S49763	gingipain R (EC 3.4.2	3.77e+02
851	6	0.7	1033	2	F71256	conserved hypotheticala	3.77e+02	924	6	A61262	collagen alpha 1(XVII	3.77e+02
852	6	0.7	1035	2	S73601	protein P200 - Mycopl	3.77e+02	925	6	A37023	IgA-specific metallo	3.77e+02
853	6	0.7	1037	2	A60153	glycoprotein Iib - ra	3.77e+02	926	6	A40891	trichyalin - sheep	3.77e+02
854	6	0.7	1037	2	S47734	probable membrane pro	3.77e+02	927	6	S76781	hypothetical protein	3.77e+02
855	6	0.7	1038	2	S38801	alpha-dextrin endo-1,	3.77e+02	928	6	T03730	antigen containing ep	3.77e+02
856	6	0.7	1039	2	I38401	ATP-driven ion pump -	3.77e+02	929	6	A65044	hypothetical protein	3.77e+02
857	6	0.7	1040	2	S50617	BEI1 protein - yeast	3.77e+02	930	6	A35140	hemolysin A precursor	3.77e+02
858	6	0.7	1047	2	A55617	masquerade precursor	3.77e+02	931	6	E70876	probable polyketidesy	3.77e+02
859	6	0.7	1053	2	B70987	probable PPE protein	3.77e+02	932	6	A28182	hemolysin A - Serrat	3.77e+02
860	6	0.7	1058	2	S50295	hypothetical protein	3.77e+02	933	6	S39059	protein BRG1 - human	3.77e+02
861	6	0.7	1061	2	D69799	cytochrome P450 / NAD	3.77e+02	934	6	T00070	hypothetical protein	3.77e+02
862	6	0.7	1091	1	ICUNL	neural cell adhesin	3.77e+02	935	6	D71630	outer membrane protei	3.77e+02
863	6	0.7	1101	2	S51823	myosin heavy chain AT	3.77e+02	936	6	JN0896	crystalline surface 1	3.77e+02
864	6	0.7	1107	2	E55226	hypothetical 123.8 kD	3.77e+02	937	6	JC4956	vitellogenin precursor	3.77e+02
865	6	0.7	1109	2	S53601	myosin-IC - human	3.77e+02	938	6	CGH04B	collagen alpha 1(IV)	3.77e+02
866	6	0.7	1114	2	I50222	deltaE1F1 - chicken	3.77e+02	939	6	T01265	starch synthase DULL1	3.77e+02
867	6	0.7	1116	2	B70476	hypothetical protein	3.77e+02	940	6	T01367	hypothetical protein	3.77e+02
868	6	0.7	1116	2	S77213	DNA-directed DNA poly	3.77e+02	941	6	S71363	probable transport pr	3.77e+02
869	6	0.7	1140	2	A48843	MHC class II transact	3.77e+02	942	6	A55426	gingipain R (EC 3.4.2	3.77e+02
870	6	0.7	1140	2	S73786	hypothetical protein	3.77e+02	943	6	S77908	hypothetical protein	3.77e+02
871	6	0.7	1146	2	S40311	integrin - fruit fly	3.77e+02	944	6	S88880	receptor DBC-205 - mo	3.77e+02
872	6	0.7	1151	2	H71347	hypothetical protein	3.77e+02	945	6	S13178	6-methylsalicylic aci	3.77e+02
873	6	0.7	1164	1	RNV28T	DNA-directed RNA poly	3.77e+02	946	6	A32230	tenascin precursor -	3.77e+02
874	6	0.7	1164	2	G36850	A24R protein - variol	3.77e+02	947	6	S35694	cut1 protein - fissio	3.77e+02
875	6	0.7	1189	2	A34817	ATPase ScII, chromoso	3.77e+02	948	6	B35049	ankyrin 1, erythrocyt	3.77e+02
876	6	0.7	1191	2	A33491	bumetanide-sensitive	3.77e+02	949	6	S01787	fatty-acid synthase (	3.77e+02
877	6	0.7	1205	2	A35547	Ca2+-transporting ATP	3.77e+02	950	6	C47521	gag-pol-like fusion p	3.77e+02
878	6	0.7	1209	2	T00373	KIAA0649 protein - hu	3.77e+02	951	6	A35049	ankyrin 1, erythrocyt	3.77e+02
879	6	0.7	1212	2	T00332	hypothetical protein	3.77e+02	952	6	SJHUK	ankyrin 1, erythrocyt	3.77e+02
880	6	0.7	1216	2	S46177	probable Ca2+-transpo	3.77e+02	953	6	A56158	eye development prote	3.77e+02
881	6	0.7	1218	2	S71376	glutamate receptor ho	3.77e+02	954	6	T02155	DNA-directed DNA poly	3.77e+02
882	6	0.7	1232	2	D64413	cobalam biosynthesi	3.77e+02	955	6	T48153	beta-myosin heavy cha	3.77e+02
883	6	0.7	1236	2	S36329	hypothetical protein	3.77e+02	956	6	S06005	myosin alpha heavy ch	3.77e+02
884	6	0.7	1242	1	DJBEC1	DNA-directed DNA poly	3.77e+02	957	6	S04090	myosin heavy chain 3,	3.77e+02
885	6	0.7	1252	2	S21178	botulinum neurotoxin	3.77e+02	958	6	S21601	hypothetical protein,	3.77e+02
886	6	0.7	1265	2	S57968	Ran-binding protein 2	3.77e+02	959	6	PN0093	genome polyprotein -	3.77e+02
887	6	0.7	1266	2	I59314	isoleucine--trNA liga	3.77e+02	960	6	S29529	genome polyprotein -	3.77e+02
888	6	0.7	1268	2	S22781	neurocan - mouse	3.77e+02	961	6	S38480	nonstructural protein	3.77e+02
889	6	0.7	1268	2	S33411	botulinum neurotoxin	3.77e+02	962	6	S60224	polyketide synthase 1	3.77e+02
890	6	0.7	1281	2	T00346	hypothetical protein	3.77e+02	963	6	T02021	nonstructural polypro	3.77e+02
891	6	0.7	1286	2	T02187	probable p-glycoprote	3.77e+02	964	6	RRWRH	genome polyprotein -	3.77e+02
892	6	0.7	1289	2	C70044	probable phosphoester	3.77e+02	965	6	S55399	genome polyprotein -	3.77e+02
893	6	0.7	1291	3	T00019	period protein homolo	3.77e+02	966	6	A56923	transcription factor	3.77e+02
894	6	0.7	1291	1	A28334	protein-tyrosine-phos	3.77e+02	967	6	A56922	transcription factor	3.77e+02
895	6	0.7	1295	1	S5ECPG	phosphoribosylformyl	3.77e+02	968	6	A38194	desmoplakin 1 - human	3.77e+02
896	6	0.7	1296	1	HNSOIF	aggregation protein a	3.77e+02	969	6	T2703	notch protein - fruit	3.77e+02
897	6	0.7	1299	2	T00261	hypothetical protein	3.77e+02	970	6	B55282	neurofibromatosis-rel	3.77e+02
898	6	0.7	1305	2	H41662	150K mating aggregate	3.77e+02	971	6	JC5196	neurofibromin 1 - rat	3.77e+02
899	6	0.7	1306	1	A31759	peptidyl-dipeptidase	3.77e+02	972	6	I54352	neurofibromin - mouse	3.77e+02

```
973 6 0.7 2871 2 A55367 fibrillin I - bovine 3.77e+02
974 6 0.7 2875 1 RRVUTW genome polyprotein - 3.77e+02
975 6 0.7 2970 1 CGHU3A collagen alpha 3(VI) 3.77e+02
976 6 0.7 3002 2 A47231 fibrillin I precursor 3.77e+02
977 6 0.7 3010 1 GNVWC genome polyprotein - 3.77e+02
978 6 0.7 3011 2 S40770 polyprotein precursor 3.77e+02
979 6 0.7 3011 1 GNVWC3 genome polyprotein - 3.77e+02
980 6 0.7 3011 1 GNVWC3 genome polyprotein - 3.77e+02
981 6 0.7 3133 2 S52093 hemocytin - silkworm 3.77e+02
982 6 0.7 3163 1 J01895 genome polyprotein - 3.77e+02
983 6 0.7 3168 2 T00296 toxin B - Escherichia 3.77e+02
984 6 0.7 3269 2 S69625 hypothetical protein 3.77e+02
985 6 0.7 3473 2 S27927 polyprotein - rice tu 3.77e+02
986 6 0.7 3649 2 S18368 delta-(L-alpha-aminoa 3.77e+02
987 6 0.7 3744 2 S46715 hypothetical protein 3.77e+02
988 6 0.7 3866 2 B48205 All-1 protein -GTE fo 3.77e+02
989 6 0.7 3869 2 A48205 All-1 protein +GTE fo 3.77e+02
990 6 0.7 4092 2 S38128 dynein heavy chain, c 3.77e+02
991 6 0.7 4302 2 A38971 polycystic kidney dis 3.77e+02
992 6 0.7 4351 2 T00252 MEGF1 protein - rat 3.77e+02
993 6 0.7 4377 2 A55575 ankryrin 3, long splic 3.77e+02
994 6 0.7 4391 2 A38096 perlecan precursor - 3.77e+02
995 6 0.7 4436 2 E71086 hypothetical protein 3.77e+02
996 6 0.7 4544 1 S02392 alpha-2-macroglobulin 3.77e+02
997 6 0.7 4545 1 S25111 alpha-2-macroglobulin 3.77e+02
998 6 0.7 4859 2 S74173 ryanodine receptor 3 3.77e+02
999 6 0.7 5147 1 IJFTFM cadherin-related tumo 3.77e+02
1000 6 0.7 15281 2 S41309 cyclosporin synthetas 3.77e+02
```

## ALIGNMENTS

```
1
RESULT 1
ENTRY JG0168 #type complete
TITLE gob-5 protein - Mouse
ORGANISM gob-5 protein Mus musculus #common_name house mouse
DATE 18-Mar-1999 #sequence_revision 18-Mar-1999 #text_change

ACCESSIONS JG0168
REFERENCE JG0168
#authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351
#title Cloning and identification of the gene gob-5, which is
#accession expressed in intestinal goblet cells in mice.
JG0168
#status preliminary
#residues 1-913 #label KOM
#cross-references DDBJ:AB016592
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 4.4%; Score 40; DB 3; Length 913;
Best Local Similarity 100.0%; Pred. No. 1.39e-107;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 615 IROGASPILRASVTALIESVNGKVTLELDNGAGADATK 654
|||||
QY 614 IROGASPILRASVTALIESVNGKVTLELDNGAGADATK 653
|||||

2
RESULT 2
ENTRY T02171 #type complete
TITLE hypothetical protein Lu-ECAM-1 - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change

ACCESSIONS T02171
REFERENCE Z14590
#authors Elble, R.C.; Widom, J.; Gruber, A.D.; Abdel-Ghany, M.;
#submission Levine, R.; Goodwin, A.; Pauli, B.U.
#description submitted to the EMBL Data Library, April 1997
#accession Cloning and characterization of Lu-ECAM-1 suggest it is an
endothelial chloride channel.
T02171
```

```

QY 154 FVHEWAHLRWG 164

RESULT 5
ENTRY GFBPT4 #type complete
TITLE gene 14 protein - phage T4
ORGANISM #formal_name phage T4
#note host Escherichia coli
DATE 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change
05-Sep-1997

ACCESSIONS JF0053; S04097
REFERENCE JF0046
#authors Selivanov, N.A.; Prilipov, A.G.; Mesyanzhinov, V.V.
#journal Nucleic Acids Res. (1989) 17:3583
#title Nucleotide sequences of bacteriophage T4 genes 13, 14 and 15.
#cross-references MUID:89263795
#accession JF0053
#status translation not shown
#molecule_type DNA
#residues 1-256 #label SEL
#cross-references EMBL:X14868; NID:g15294; PID:g15296
COMMENT Gene 14 protein is a component of the phage neck.
GENETICS
#gene 14
#map_position 93.55-94.25
CLASSIFICATION #superfamily phage T4 gene 14 protein
SUMMARY #length 256 #molecular_weight 29573 #checksum 2507

Query Match 0.9%; Score 8; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 6.56e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 43 ADVLVAES 50
|||||||
QY 102 ADLVAES 109

RESULT 6
ENTRY S66279 #type complete
TITLE proline dehydrogenase (EC 1.5.99.8) /
1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12) -
Salmonella typhimurium
CONTAINS 1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12);
proline dehydrogenase (EC 1.5.99.8)
ORGANISM #formal_name Salmonella typhimurium
DATE 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
17-Mar-1999
ACCESSIONS S66279; S33716; B39192; S03817; S31910
REFERENCE S66279
#authors Maloy, S.R.
#submission submitted to the EMBL Data Library, April 1994
#accession S66279
#molecule_type DNA
#residues 1-1320 #label MAL
#cross-references EMBL:X70843; NID:g470179; PID:g470180
#experimental_source strain Ir2
#note this is a revision to the sequence from reference S33716
REFERENCE S33716
#authors Allen, S.W.; Senti-Willis, A.; Maloy, S.R.
#journal Nucleic Acids Res. (1993) 21:1676
#title DNA sequence of the putA gene from Salmonella typhimurium: a
bifunctional membrane-associated dehydrogenase that binds
DNA.
#cross-references MUID:93241961
#accession S33716
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-19, 'WPRRVVIA', 28-272, 'ENWRK', 278-279, 'SAILTICWAN',
290-368, 'APKRRIVWRSICWKSASNPWRAGTC', 398-592,
'RKVRLAYIRIKRRCRAICTA', 613-907, 'CFQRRTTLFR', 919-1034,
'DV', 1037-1152,
'RRYRSADNSPIWPRPARSACYGRPASAIPOGCCFVNGYAWLWMNRTR'
#label ALL

##cross-references EMBL:X70843
##note the nucleotide sequence was submitted to the EMBL Data
Library, January 1993
##note this sequence has been revised in reference S66279

REFERENCE S53664
#authors Ling, M.; Allen, S.W.; Wood, J.M.
#journal J. Mol. Biol. (1994) 243:950-956
#title Sequence analysis identifies the proline dehydrogenase and
Delta(1)-pyrroline-5-carboxylate dehydrogenase domains of
the multifunctional Escherichia coli PutA protein.
#cross-references MUID:95055736
#contents annotation
REFERENCE A39192
#authors Ostrovsky de Spicer, P.; O'Brien, K.; Maloy, S.
#journal J. Bacteriol. (1991) 173:211-219
#title Regulation of proline utilization in Salmonella typhimurium:
a membrane-associated dehydrogenase binds DNA in vitro.
#cross-references MUID:91100285
#accession B39192
#status preliminary
#molecule_type DNA
#residues 1-26 #label OST
##note the authors translated the codon CGT for residue 24 as
Gly

REFERENCE S03816
#authors Hahn, D.R.; Myers, R.S.; Kent, C.R.; Maloy, S.R.
#journal Mol. Gen. Genet. (1988) 213:125-133
#title Regulation of proline utilization in Salmonella typhimurium:
molecular characterization of the put operon, and DNA
sequence of the put control region.
#cross-references MUID:89127131
#accession S03817
#molecule_type DNA
#residues 1-13, 'HARTDQ', 20-26 #label HAH
##cross-references EMBL:X12569

GENETICS
#gene putA
#map_position 22 min
FUNCTION
#description transfers electrons from proline to the respiratory chain;
catalyzes proline oxidation to 1-pyrroline-5-carboxylate
#pathway proline utilization
#note membrane-bound with proline
FUNCTION <HPC>
#description catalyzes hydrolysis of 1-pyrroline-5-carboxylate to
gamma-glutamyl semialdehyde
#pathway proline utilization
FUNCTION <RPA>
#description transcriptional repressor controls expression of genes putP
and putA in response to proline supply
CLASSIFICATION #superfamily bifunctional protein putA
KEYWORDS DNA binding; FAD; flavoprotein; membrane-associated protein;
NAD; oxidoreductase; transcription regulation
FEATURE
883 #active_site Glu #status predicted\
917 #active_site Cys #status predicted\
SUMMARY #length 1320 #molecular_weight 144187 #checksum 3619

Query Match 0.9%; Score 8; DB 2; Length 1320;
Best Local Similarity 100.0%; Pred. No. 6.56e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1228 LAKRLPAA 1235
|||||||
QY 371 LAKRLPAA 378

RESULT 7
ENTRY D71614 #type complete
TITLE hypothetical protein PF0460c - malaria parasite (Plasmodium
falciparum)
ORGANISM #formal_name Plasmodium falciparum
DATE 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change

```

```

13-Nov-1998
ACCESSIONS
REFERENCE
#authors
Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.;
Aravind, L.; Koonin, E.V.; Shalom, S.; Mason, T.; Yu, K.;
Fuji, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.;
Lai, Z.; Schwartz, D.C.; Pertea, M.; Salzberg, S.; Zhou,
L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.L.
#journal
Science (1998) 282:1136-1132
#title
Chromosome 2 sequence of the human malaria parasite
Plasmodium falciparum.
#accession
D71614
#status
preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type
DNA
##residues
1-2573 ##label
GAR
##cross-references
GB:AE001396; GB:AE001362; NID:g3845188; PID:g3845190;
TIGR:PF04600
##experimental_source
clone 3D7
GENETICS
#gene
PF04600
SUMMARY
#length
2573 #molecular_weight
308142 #checksum
4299
Query Match 0.9%; Score 8; DB 2; Length 2573;
Best Local Similarity 100.0%; Pred. No. 6.56e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 2211 EYNDKX 2218
|||||||
Qy 168 EYNDKX 175

RESULT 8
ENTRY
TITLE
transferrin binding protein 2 - Neisseria meningitidis
(fragment)
ORGANISM
#formal_name
Neisseria meningitidis
strain 528
#variety
27-Apr-1996 #sequence_revision
13-Mar-1997 #text_change
17-Mar-1999
ACCESSIONS
S61547
REFERENCE
S61544
#authors
Rokbi, B.; Waire-Wilmotte, G.; Mazarin, V.; Fourrignon, L.;
Lissolo, L.; Quentin-Millet, M.J.
#journal
FEMS Microbiol. Lett. (1995) 132:277-283
#title
Variable sequences in a mosaic-like domain of meningococcal
tbp2 encode immunoreactive epitopes.
#cross-references
M01D:96039602
#accession
S61547
#status
preliminary; translation not shown
##molecule_type
DNA
##residues
1-35 ##label
ROK
##cross-references
EMBL:X88868
CLASSIFICATION
#superfamily
bacterial pathogen transferrin-binding protein;
tonB-dependent receptor amino-terminal homology;
tonB-dependent receptor carboxyl-terminal homology
KEYWORDS
membrane protein
SUMMARY
#length
35 #checksum
6464
Query Match 0.8%; Score 7; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 7.01e-00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 11 AAASGGT 17
|||||||
Qy 377 AAASGGT 383

RESULT 9
ENTRY
TITLE
trypsin (EC 3.4.21.4) A precursor - marbled lungfish
(fragment)

```



```

REFERENCE
#authors
Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;
Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo,
S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
Yasuda, M.; Tabata, S.
#journal
DNA Res. (1996) 3:109-136
#title
Sequence analysis of the genome of the unicellular
cyanobacterium Synechocystis sp. PCC6803. II. Sequence
determination of the entire genome and assignment of
potential protein-coding regions.
#cross-references MUID:97061201
#accession
S75106
#status preliminary
#molecule_type DNA
#residues 1-80 #label KAN
#cross-references EMBL:D90910; GB:AB001339; NID:g1652956; PID:d1018701;
PID:g1653051
#note
the nucleotide sequence was submitted to the EMBL Data
Library, June 1996
#length 80 #molecular-weight 9170 #checksum 3659
SUMMARY
Query Match 0.8%; Score 7; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.01e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 27 KFNSLQ 33
QY 802 KFNSLQ 808
|||||

RESULT 12
ENTRY
H70074 #type complete
TITLE
hypoetical protein yxeE - Bacillus subtilis
ORGANISM
Bacillus subtilis
DATE
05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
24-Sep-1998
ACCESSIONS
A69580
REFERENCE
#authors
Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;
Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
Choi, S.K.; Codani, J.J.; Conner, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrarri, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itava, M.; Jones, L.; Joris, B.; Karamata, D.;
Kashara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
I.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetle, D.; Porwolik, S.; Prescott,
A.M.; Pressecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.J.; Serron, P.; Shin, B.S.; Soldo,
B.; Sorokin, A.; Taccioni, E.; Takagi, T.; Tanahashi, H.;
Takemaru, K.; Takeuchi, M.; Takakoshi, A.; Tanaka, T.;
Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
Wambutt, R.; Wedler, E.; Wedler, H.; Weltzenegger, T.;
Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,

```

---

```

K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.;
Yoshikawa, H.; Danchin, A.
#journal
Nature (1997) 390:249-256
#title
The complete genome sequence of the Gram-positive bacterium
Bacillus subtilis.
#cross-references MUID:98044033
#accession
G70074
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-121 #label KUN
#cross-references GB:Z99124; GB:AL009126; NID:g2636442; PID:el184683;
PID:g2636504
#experimental_source strain 168
GENETICS
#gene
yxeE
SUMMARY
#length 121 #molecular-weight 14714 #checksum 9005
Query Match 0.8%; Score 7; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 7.01e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 RLNLNQ 101
QY 322 RLNLNQ 328
|||||

RESULT 13
ENTRY
H70435 #type complete
TITLE
hypoetical protein aq_1570 - Aquifex aeolicus
ORGANISM
Aquifex aeolicus
DATE
08-May-1998 #sequence_revision 08-May-1998 #text_change
08-May-1998
ACCESSIONS
H70435
REFERENCE
#authors
Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.;
Lenox, A.L.; Graham, D.E.; Overbeek, R.; Sneed, M.A.;
Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short,
J.M.; Olson, G.J.; Swanson, R.V.
#journal
Nature (1998) 392:353-358
#title
The complete genome of the hyperthermophilic bacterium
Aquifex aeolicus.
#cross-references MUID:98196666
#accession
H70435
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-128 #label AQF
#cross-references GB:AE000746; NID:g2983925; PID:g2983937; GB:AE000657
#experimental_source strain Vfs5
GENETICS
#gene
aq_1570
SUMMARY
#length 128 #molecular-weight 15365 #checksum 7819
Query Match 0.8%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 7.01e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 78 LEELSKM 84
QY 445 LEELSKM 451
|||||

RESULT 14
ENTRY
G70646 #type complete
TITLE
hypoetical protein RV3142c - Mycobacterium tuberculosis
(strain H37RV)
ORGANISM
Mycobacterium tuberculosis
DATE
17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
17-Jul-1998
ACCESSIONS
G70646
REFERENCE
#authors
Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,

```

C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry  
 III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;  
 Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;  
 Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
 Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;  
 Murphy, L.; Oliver, S.; Osborne, J.; Ouail, M.A.;  
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;  
 Skelton, S.; Squares, S.; Squires, R.; Sulston, J.E.;  
 Taylor, K.; Whitehead, S.; Barrell, B.G.

#journal Nature (1998) 393:537-544  
 #title Deciphering the biology of Mycobacterium tuberculosis from  
 the complete genome sequence.

#cross-references MUID:98295987

#accession G70646

##status preliminary; nucleic acid sequence not shown;

##molecule\_type DNA translation not shown

##residues 1-142 #label COL

##cross-references GB:283867; GB:AL123456; NID:g3261695; PID:e290946;

PID:g1781225

##experimental\_source strain H37Rv

# GENETICS

#gene Rv3142C

#SUMMARY #length 142 #molecular-weight 15935 #checksum 1401

Query Match 0.8%; Score 7; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 7.01e+00;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 VVAIDPN 61

QY 36 VVAIDPN 42

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Search completed: Sat Aug 28 14:34:46 1999

, Job time : 77 secs.

\*\*\*\*\*

W P E R E H

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:30:11 1999; MasPar time 24.54 Seconds  
Tabular output not generated. 1053.033 Million cell updates/sec

Title: >US-09-049-696-41  
Description: (1-914) from US09049696.pgp  
Perfect Score: 914  
Sequence: 1 MGPKSSVFIILHLLEGAL.....GIHLKIMKMWIGELQLSIA 914

Scoring table: TABLE unitprotatable  
Gap 60

Searched: 77977 seqs, 28258293 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 4.053; Variance 0.386; scale 10.502

pred, No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	14	1.5	903	1	ECLC_BOVIN EPITHELIAL CHLORIDE CH	6.75e-20
2	8	0.9	256	1	VG14_BP14 NECK PROTEIN GP14	1.73e-02
3	8	0.9	300	1	Y4TE_RH1SN PROBABLE AMINO-ACID AB	1.73e-02
4	8	0.9	1320	1	PUTA_SALT1 PROLINE DEHYDROGENASE	1.73e-02
5	7	0.8	121	1	YXEE_BACSU HYPOTHETICAL 14.7 KD P	2.80e+00
6	7	0.8	139	1	RPB6_ASF2 DNA-DIRECTED RNA POLYM	2.80e+00
7	7	0.8	147	1	RPB6_ASF7 T-LYMPHOCYTE MATURATION	2.80e+00
8	7	0.8	153	1	MAL_RAT T-LYMPHOCYTE MATURATION	2.80e+00
9	7	0.8	153	1	MAL_CANFA T-LYMPHOCYTE MATURATION	2.80e+00
10	7	0.8	153	1	YR35_MYCTU HYPOTHETICAL 16.4 KD P	2.80e+00
11	7	0.8	154	1	BCSD_ACEXY CELLULOSE SYNTHASE OPE	2.80e+00
12	7	0.8	156	1	PRIA_LYCES PATHOGENESIS-RELATED P	2.80e+00
13	7	0.8	175	1	E320_ADE03 EARLY E3 20.1 KD GLYCO	2.80e+00
14	7	0.8	179	1	E320_ADE03 EARLY E3 20.1 KD GLYCO	2.80e+00
15	7	0.8	194	1	Y0Y8_CAEEL HYPOTHETICAL 23.0 KD P	2.80e+00
16	7	0.8	201	1	SYA_RH1LV ALANIL-TRNA SYNTHETASE	2.80e+00
17	7	0.8	211	1	UPP_LACLA URACIL PHOSPHORIBOSYL	2.80e+00
18	7	0.8	243	1	CD48_HUMAN B-LYMPHOCYTE ACTIVATIO	2.80e+00
19	7	0.8	254	1	Y375_METJA HYPOTHETICAL PROTEIN M	2.80e+00
20	7	0.8	270	1	YC87_METJA HYPOTHETICAL PROTEIN M	2.80e+00
21	7	0.8	297	1	ARCC_ECOLI CARBANATE KINASE (EC 2	2.80e+00
22	7	0.8	299	1	YDEX_HAEIN HYPOTHETICAL PROTEIN H	2.80e+00
23	7	0.8	302	1	CANI_RABIT CALPAIN 1, LARGE [CATA	2.80e+00

1	GAG_MSVER	GAG POLYPROTEIN [CONTA	2.80e+00
1	YFM4_MICOL	HYPOTHETICAL PROTEIN I	2.80e+00
1	MIIP_SCMV1	MT-1 PROTOXIN PRECURSO	2.80e+00
1	Y050_SYNY1	HYPOTHETICAL 36.1 KD P	2.80e+00
1	HXAL_HUMAN	HOMEOBOX PROTEIN HOX-A	2.80e+00
1	Y4TR_RH1SN	PROBABLE PEPTIDE ABC T	2.80e+00
1	PEXE_PICAN	PEROXISOMAL MEMBRANE P	2.80e+00
1	VM17_BORHE	VARIABLE MAJOR OUTER M	2.80e+00
1	RIR2_ECOLI	RIBONUCLEOSIDE-DIPHOSP	2.80e+00
1	NG10_HUMAN	MHC CLASS I NK CELL RE	2.80e+00
1	YKJQ_ECOLI	HYPOTHETICAL 43.4 KD P	2.80e+00
1	DYNC_HUMAN	DYNACTIN, 50 KD ISOFOR	2.80e+00
1	SVS2_RAT	SEMINAL VESICLE SECRET	2.80e+00
1	YEBA_ECOLI	HYPOTHETICAL 46.7 KD P	2.80e+00
1	GAG_MSMTV	GAG POLYPROTEIN [CONTA	2.80e+00
1	ST11_SCHPO	TRANSCRIPTION FACTOR S	2.80e+00
1	TRPE_ACICA	ANTHRANILATE SYNTHASE	2.80e+00
1	GAG_MLVDU	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVDE	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVCB	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVVD	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVAV	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVBM	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVFF	GAG POLYPROTEIN [CORE	2.80e+00
1	GAG_MVMO	GAG POLYPROTEIN R65 [C	2.80e+00
1	GAG_MLVFP	GAG POLYPROTEIN [CONTA	2.80e+00
1	GAG_MLVF5	GAG POLYPROTEIN [CORE	2.80e+00
1	GAG_MLVH0	GAG POLYPROTEIN [CONTA	2.80e+00
1	HEMA_IADH7	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH1	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH4	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IAGHK	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH3	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH6	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH3	HEMAGGLUTININ PRECURSO	2.80e+00
1	HEMA_IADH2	HEMAGGLUTININ PRECURSO	2.80e+00
1	CC14_YEAST	PROBABLE PROTEIN-TYROS	2.80e+00
1	HEMA_IAMAO	HEMAGGLUTININ PRECURSO	2.80e+00
1	IFP2_MEITH	TRANSLATION INITIATION	2.80e+00
1	GHR_COLLI	GROWTH HORMONE RECEPT	2.80e+00
1	PLB1_PENNO	LYSOPHOSPHOLIPASE PREC	2.80e+00
1	YEA7_YEAST	HYPOTHETICAL 72.5 KD P	2.80e+00
1	HMOC_DROME	HOMEOIC PROTEIN ORTHO	2.80e+00
1	ANX6_HUMAN	ANNEXIN VI (LIPOCORTIN	2.80e+00
1	ANX6_RAT	ANNEXIN VI (LIPOCORTIN	2.80e+00
1	ANX6_MOUSE	ANNEXIN VI (LIPOCORTIN	2.80e+00
1	CAN1_HUMAN	CALPAIN 1, LARGE [CATA	2.80e+00
1	DNLJ_TREPA	DNA LIGASE (EC 6.5.1.2	2.80e+00
1	FIBA_HUMAN	FIBRINOGEN ALPHA/ALPHA	2.80e+00
1	POLG_ENMG3	GENOME POLYPROTEIN [CO	2.80e+00
1	TSF4_RAT	THROMBOSPONDIN 4 PRECU	2.80e+00
1	DSG3_HUMAN	DESMOGLEIN 3 PRECURSOR	2.80e+00
1	DDX9_CAEEL	PROBABLE ATP-DEPENDENT	2.80e+00
1	TR12_HUMAN	THYROID RECEPTOR INTER	2.80e+00
1	POLG_HRV2	GENOME POLYPROTEIN [CO	2.80e+00
1	POLG_EMCV	GENOME POLYPROTEIN [CO	2.80e+00
1	POLG_EMCVD	GENOME POLYPROTEIN [CO	2.80e+00
1	POLG_EMCVB	GENOME POLYPROTEIN [CO	2.80e+00
1	PCP7_CANPG	CUTICLE PROTEIN CP459	2.07e+02
1	E1A6_ADE07	EARLY E1A 6.3 KD PROTE	2.07e+02
1	ROP_ECOLI	REGULATORY OF G-PROTEIN	2.07e+02
1	RGSA_RAT	REGULATOR OF G-PROTEIN	2.07e+02
1	RGSS_RAT	REGULATOR OF G-PROTEIN	2.07e+02
1	ATP8_GORGO	ATP SYNTHASE PROTEIN 8	2.07e+02
1	ATP8_PANPA	ATP SYNTHASE PROTEIN 8	2.07e+02
1	ATP8_HUMAN	ATP SYNTHASE PROTEIN 8	2.07e+02
1	YF77_HAEIN	HYPOTHETICAL PROTEIN H	2.07e+02
1	RL29_AQUAE	50S_RIBOSOMAL PROTEIN	2.07e+02

97	6	0.7	84	1	VEA_HPVI	PROBABLE E4 PROTEIN.	2.07e+02	170	6	0.7	183	1	SAV1_STRVL	STREPTAVIDIN V1 PRECUR	2.07e+02
98	6	0.7	85	1	YPBS_BACSU	HYPOTHETICAL 9.9 KD PR	2.07e+02	171	6	0.7	184	1	HXB1_AMBME	HOMEOBOX PROTEIN HOX-B	2.07e+02
99	6	0.7	91	1	MERP_SHEPU	MERCURIC TRANSPORT PRO	2.07e+02	172	6	0.7	185	1	SCP1_BRALA	SARCOPLASMIC CALCULIN-B	2.07e+02
100	6	0.7	91	1	DBH_CLOPA	DNA-BINDING PROTEIN HU	2.07e+02	173	6	0.7	191	1	GDIR_CAEEL	PROBABLE RHO GDP-DISSO	2.07e+02
101	6	0.7	92	1	CYK5_RAT	CYTOCHROME B5, OUTER M	2.07e+02	174	6	0.7	192	1	Y3_SOCMV	HYPOTHETICAL PROTEIN 3	2.07e+02
102	6	0.7	95	1	HX08_HUMAN	HOMEOBOX PROTEIN HOX-D	2.07e+02	175	6	0.7	193	1	FLA5_HALHA	FLAGELLIN B3 PRECURSOR	2.07e+02
103	6	0.7	100	1	XN08_HAEIN	HYPOTHETICAL PROTEIN H	2.07e+02	176	6	0.7	194	1	YG13_BUCAP	HYPOTHETICAL 23.2 KD P	2.07e+02
104	6	0.7	105	1	YGRM_MICPU	HYPOTHETICAL PROTEIN I	2.07e+02	177	6	0.7	194	1	TNP0_ECOLI	TRANSPOSON TN2501 RESO	2.07e+02
105	6	0.7	105	1	THIQ_RHOSH	THIOREDOXIN (TRX) .	2.07e+02	178	6	0.7	196	1	CH19_DROGR	CHORION PROTEIN S19	2.07e+02
106	6	0.7	106	1	CYAY_ECOLI	CVAY PROTEIN.	2.07e+02	179	6	0.7	196	1	FLA4_HALHA	FLAGELLIN B2 PRECURSOR	2.07e+02
107	6	0.7	109	1	V42_BPT3	GENE 4.2 PROTEIN.	2.07e+02	180	6	0.7	196	1	HP20_TAMAS	HIBERNATION-ASSOCIATED	2.07e+02
108	6	0.7	109	1	FLIE_PSEAE	FLAGELLAR HOOK-BASAL B	2.07e+02	181	6	0.7	196	1	FLA1_HALHA	FLAGELLIN A1 PRECURSOR	2.07e+02
109	6	0.7	113	1	YF65_SYNY3	HYPOTHETICAL 12.5 KD P	2.07e+02	182	6	0.7	198	1	TRPG_LACLA	ANTHRANILATE SYNTHASE	2.07e+02
110	6	0.7	113	1	GTHB_MURCI	GONADOTROPIN BETA CHAI	2.07e+02	183	6	0.7	203	1	RB30_HUMAN	RAS-RELATED PROTEIN RA	2.07e+02
111	6	0.7	113	1	RLA2_AITAL	60S ACIDIC RIBOSOMAL P	2.07e+02	184	6	0.7	203	1	HPR_BACSU	PROTEASE PRODUCTION RE	2.07e+02
112	6	0.7	113	1	SMPA_ECOLI	SMALL PROTEIN A.	2.07e+02	185	6	0.7	205	1	FLRE_BOVIN	FLAVIN REDUCTASE (EC 1	2.07e+02
113	6	0.7	116	1	SP21_BACCO	ANTI-SIGMA F FACTOR AN	2.07e+02	186	6	0.7	205	1	FKBB_ECOLI	FKBP-TYPE 22 KD PEPTID	2.07e+02
114	6	0.7	118	1	CY31_DESDN	CYTOCHROME C3, 13 KD	2.07e+02	187	6	0.7	208	1	HIS1_LACLA	ATP PHOSPHORIBOSYLTRAN	2.07e+02
115	6	0.7	119	1	YRAN_HAEIN	HYPOTHETICAL PROTEIN H	2.07e+02	188	6	0.7	208	1	GTP_CAEEL	GLUTATHIONE S-TRANSFER	2.07e+02
116	6	0.7	123	1	GALA_HUMAN	GALANIN PRECURSOR.	2.07e+02	189	6	0.7	208	1	UPP_BACCL	URACIL PHOSPHORIBOSYLT	2.07e+02
117	6	0.7	124	1	VC11_PEA	VICILIN, 14 KD COMPOSE	2.07e+02	190	6	0.7	208	1	KGUA_MYCTU	GUANYLATE KINASE (EC 2	2.07e+02
118	6	0.7	124	1	VATF_DROME	VACUOLAR ATP SYNTHASE	2.07e+02	191	6	0.7	209	1	YALA_SCHPO	HYPOTHETICAL 23.8 KD P	2.07e+02
119	6	0.7	124	1	VATF_MANSE	VACUOLAR ATP SYNTHASE	2.07e+02	192	6	0.7	209	1	NUGM_PARPR	PROBABLE NADH-UBIQUINO	2.07e+02
120	6	0.7	125	1	RNP_SPAEH	RIBONUCLEASE PANCREATI	2.07e+02	193	6	0.7	209	1	RYL2_YARLI	RAS-LIKE GTP-BINDING P	2.07e+02
121	6	0.7	125	1	BOTE_BOTJA	BOTROCETIN, BETA CHAIN	2.07e+02	194	6	0.7	211	1	RG52_MOUSE	REGULATOR OF G-PROTEIN	2.07e+02
122	6	0.7	127	1	VATF_ANOGA	VACUOLAR ATP SYNTHASE	2.07e+02	195	6	0.7	211	1	RG52_HUMAN	REGULATOR OF G-PROTEIN	2.07e+02
123	6	0.7	132	1	COAT_BPSP	COAT PROTEIN.	2.07e+02	196	6	0.7	211	1	UL92_HSV7J	PROTEIN U63 .	2.07e+02
124	6	0.7	132	1	SUCA_HUMAN	SUCCINYL-COA LIGASE [G	2.07e+02	197	6	0.7	213	1	GPH_AQUAE	PHOSPHOGLYCULATE PHOSP	2.07e+02
125	6	0.7	133	1	NL22_PARJU	PROBABLE NONSPECIFIC L	2.07e+02	198	6	0.7	213	1	ER21_CAEEL	PUTATIVE ER LUMEN PROT	2.07e+02
126	6	0.7	133	1	NL21_PARJU	PROBABLE NONSPECIFIC L	2.07e+02	199	6	0.7	213	1	HS49_YEAST	HSH49 PROTEIN.	2.07e+02
127	6	0.7	134	1	YQJE_ECOLI	HYPOTHETICAL 15.1 KD P	2.07e+02	200	6	0.7	215	1	G25L_CANFA	GLYCOPROTEIN 25L PRECU	2.07e+02
128	6	0.7	140	1	HSJL_ECOLI	HEAT SHOCK PROTEIN HSL	2.07e+02	201	6	0.7	215	1	VM12_YEAST	VACUOLAR ATPASE ASSEMB	2.07e+02
129	6	0.7	141	1	RL23_TRYCR	60S RIBOSOMAL PROTEIN	2.07e+02	202	6	0.7	217	1	YAAF_BACSU	HYPOTHETICAL 25.4 KD P	2.07e+02
130	6	0.7	142	1	TOLR_ECOLI	TOLR PROTEIN.	2.07e+02	203	6	0.7	225	1	UREF_YERPS	UREASE ACCESSORY PROTE	2.07e+02
131	6	0.7	146	1	HBD_SAISC	HEMOGLOBIN DELTA CHAIN	2.07e+02	204	6	0.7	226	1	BASP_BOVIN	BASP1 PROTEIN.	2.07e+02
132	6	0.7	146	1	HBD_SAGMY	HEMOGLOBIN DELTA CHAIN	2.07e+02	205	6	0.7	228	1	UREF_YEREN	UREASE ACCESSORY PROTE	2.07e+02
133	6	0.7	147	1	PTVA_ECOLI	PTS SYSTEM, MANNITOL (	2.07e+02	206	6	0.7	231	1	TRYP_PIG	TRYPSIN PRECURSOR (EC	2.07e+02
134	6	0.7	148	1	CA19_MOUSE	COLLAGEN ALPHA 1(I)X) C	2.07e+02	207	6	0.7	232	1	YFJ0_YEAST	HYPOTHETICAL 25.7 KD P	2.07e+02
135	6	0.7	149	1	DH15_WHEAT	DEHYDRIN RAB 15.	2.07e+02	208	6	0.7	234	1	CD3L_HUMAN	CD30 LIGAND (CD30-L) (	2.07e+02
136	6	0.7	149	1	ENRN_BPT7	ENDOXYRIBONUCLEASE	2.07e+02	209	6	0.7	234	1	RPE_METJA	RIBULOSE-PHOSPHATE 3-E	2.07e+02
137	6	0.7	150	1	YPS9_YEAST	HYPOTHETICAL 16.9 KD P	2.07e+02	210	6	0.7	236	1	CUTF_ECOLI	COPPER HOMEOSTASIS PRO	2.07e+02
138	6	0.7	151	1	YKA2_YEAST	HYPOTHETICAL 16.7 KD P	2.07e+02	211	6	0.7	237	1	YGH5_ECOLI	HYPOTHETICAL 26.3 KD A	2.07e+02
139	6	0.7	151	1	HBP_WHEAT	POTENTIAL HEME-BINDING	2.07e+02	212	6	0.7	238	1	YRN7_CAEEL	HYPOTHETICAL 25.0 KD P	2.07e+02
140	6	0.7	152	1	ENRN_BPT3	ENDOXYRIBONUCLEASE	2.07e+02	213	6	0.7	243	1	TRY1_BOVIN	TRYPSINOGEN, CATIONIC	2.07e+02
141	6	0.7	154	1	AR16_YEAST	ARP2/3 COMPLEX 16 KD S	2.07e+02	214	6	0.7	243	1	HXB8_MOUSE	HOMEOBOX PROTEIN HOX-B	2.07e+02
142	6	0.7	155	1	YBBA_ECOLI	HYPOTHETICAL 17.3 KD P	2.07e+02	215	6	0.7	244	1	YJ24_YEAST	HYPOTHETICAL 27.4 KD P	2.07e+02
143	6	0.7	157	1	FMP1_PSEAE	FIMBRIAL PROTEIN PRECU	2.07e+02	216	6	0.7	244	1	TRY2_XENLA	TRYPSINOGEN PRECURSOR.	2.07e+02
144	6	0.7	158	1	PRI_ASPOF	PATHOGENESIS-RELATED P	2.07e+02	217	6	0.7	245	1	CTRA_BOVIN	CHYMOTRYPSINOGEN A (EC	2.07e+02
145	6	0.7	159	1	YE19_SYNY3	HYPOTHETICAL 17.7 KD P	2.07e+02	218	6	0.7	245	1	CTRB_GADMO	CHYMOTRYPSIN B (EC 3.4	2.07e+02
146	6	0.7	160	1	NUGM_PARLI	NADH-UBIQUINONE OXIDOR	2.07e+02	219	6	0.7	246	1	TRY2_RAT	TRYPSINOGEN II, ANIONI	2.07e+02
147	6	0.7	161	1	XKDI_BACSU	PHAGE-LIKE ELEMENT PBS	2.07e+02	220	6	0.7	246	1	TRY1_CANFA	TRYPSINOGEN I, ANIONIC	2.07e+02
148	6	0.7	162	1	YGW4_YEAST	HYPOTHETICAL 18.1 KD P	2.07e+02	221	6	0.7	246	1	TRY1_CANFA	TRYPSINOGEN I, ANIONIC	2.07e+02
149	6	0.7	164	1	Y400_SYNY3	HYPOTHETICAL 18.3 KD P	2.07e+02	222	6	0.7	246	1	TRYP_MOUSE	PROCESSING AND TRANSPOR	2.07e+02
150	6	0.7	165	1	MLRV_RAT	MYOSIN REGULATORY LIGH	2.07e+02	223	6	0.7	247	1	YQFQ_BACSU	HYPOTHETICAL 25.9 KD P	2.07e+02
151	6	0.7	165	1	MLRV_MOUSE	MYOSIN REGULATORY LIGH	2.07e+02	224	6	0.7	247	1	TRY2_BOVIN	TRYPSINOGEN, ANIONIC P	2.07e+02
152	6	0.7	168	1	TRY3_LUCCU	TRYPSIN ALPHA-3 (EC 3.	2.07e+02	225	6	0.7	247	1	TRY1_HUMAN	TRYPSINOGEN I PRECURSO	2.07e+02
153	6	0.7	168	1	ATPD_HUMAN	ATP SYNTHASE DELTA CHA	2.07e+02	226	6	0.7	247	1	TRY3_RAT	TRYPSINOGEN III, CATIO	2.07e+02
154	6	0.7	172	1	UL45_HSV1K	PROTEIN UL45 (18 KD PR	2.07e+02	227	6	0.7	247	1	TRY3_RAT	TRYPSINOGEN III, CATIO	2.07e+02
155	6	0.7	172	1	UL45_HSV1M	PROTEIN UL45 (18 KD PR	2.07e+02	228	6	0.7	248	1	PSPA_HUMAN	PULMONARY SURFACTANT-A	2.07e+02
156	6	0.7	172	1	UL45_HSV1L	PROTEIN UL45.	2.07e+02	229	6	0.7	248	1	TRY3_CHICK	TRYPSINOGEN II-P29 PRE	2.07e+02
157	6	0.7	174	1	MOBB_ECOLI	MOLYBDOPTEIN-GUANINE	2.07e+02	230	6	0.7	248	1	NKPI_RAT	NATURAL KILLER CELL PR	2.07e+02
158	6	0.7	175	1	YFGB_YEAST	HYPOTHETICAL 16.5 KD P	2.07e+02	231	6	0.7	251	1	HIS6_PSEAE	HISF PROTEIN (CYCLASE)	2.07e+02
159	6	0.7	175	1	SFAG_ECOLI	S FIMBRIA ADHESIN PROT	2.07e+02	232	6	0.7	252	1	CLCD_RHOOP	CARBOXYMETHYLENEBUTENO	2.07e+02
160	6	0.7	175	1	IF3_AQUAE	TRANSLATION INITIATION	2.07e+02	233	6	0.7	252	1	HIS6_RHOSH	HISF PROTEIN (CYCLASE)	2.07e+02
161	6	0.7	175	1	CBUP_RAT	CALCINEURIN B-LIKE PRO	2.07e+02	234	6	0.7	252	1	HIS6_RHOCA	HISF PROTEIN (CYCLASE)	2.07e+02
162	6	0.7	176	1	YD2F_SCHPO	VERY HYPOTHETICAL 20.6	2.07e+02	235	6	0.7	253	1	PDH1_CAEEL	PUTATIVE PYRIDOXAMINE	2.07e+02
163	6	0.7	177	1	BRO_DROME	BROTHER PROTEIN.	2.07e+02	236	6	0.7	253	1	HIS6_AQUAE	HISF PROTEIN (CYCLASE)	2.07e+02
164	6	0.7	177	1	RM05_ACACA	MITOCHONDRIAL 60S RIBO	2.07e+02	237	6	0.7	253	1	GBL1_TYLHE	GIANT EXTRACELLULAR HE	2.07e+02
165	6	0.7	180	1	YAF5_YEAST	HYPOTHETICAL 19.9 KD P	2.07e+02	238	6	0.7	255	1	TRY4_LUCCU	TRYPSIN ALPHA-4 PRECUR	2.07e+02
166	6	0.7	180	1	YPKW_THECU	HYPOTHETICAL PROTEIN I	2.07e+02	239	6	0.7	256	1	YMS4_MAIZE	HYPOTHETICAL 29 KD PRO	2.07e+02
167	6	0.7	182	1	YOPO_YEREN	YOPO PROTEIN PRECURSOR	2.07e+02	240	6	0.7	256	1	LPI_BOMBO	LOW MOLECULAR LIPOPROT	2.07e+02
168	6	0.7	183	1	SAV_STRAV	STREPTAVIDIN PRECURSOR	2.07e+02	241	6	0.7	259	1	1430_ARATH	14-3-3-LIKE PROTEIN GF	2.07e+02
169	6	0.7	183	1	SAV2_STRVL	STREPTAVIDIN V2 PRECUR	2.07e+02	242	6	0.7	259	1	RFA2_KLEPN	O-ANTIGEN EXPORT SYSTE	2.07e+02

243	1	PKSH_BACSU	PUTATIVE POLYPEPTIDE BI	2.07e+02	316	1	YOH1_ECOLI	HYPOTHETICAL 35.2 KD P	2.07e+02
244	1	RPOD_HALMA	DNA-DIRECTED RNA POLYM	2.07e+02	317	1	CAIC_RAT	COLLAGEN ALPHA 1(XII)	2.07e+02
245	1	MTM2_MORBO	MODIFICATION METHYLASE	2.07e+02	318	1	KLAO_SOYBN	60S ACIDIC RIBOSOMAL P	2.07e+02
246	1	GTTL_CHICK	GLUTATHIONE S-TRANSFER	2.07e+02	319	1	FCEE2_HUMAN	LOW AFFINITY IMMUNOGLO	2.07e+02
247	1	HIS6_SYNY3	HISF PROTEIN (CYCLASE)	2.07e+02	320	1	RM10_YEAST	MITOCHONDRIAL 60S RIBO	2.07e+02
248	1	COBK_SYNY3	PRECORRIN-6X REDUCTASE	2.07e+02	321	1	GCA_RAT	IG GAMMA-2A CHAIN C RE	2.07e+02
249	1	HIS6_AZOB	HISF PROTEIN (CYCLASE)	2.07e+02	322	1	QOX2_BACSU	QUINOL OXIDASE POLYPEP	2.07e+02
250	1	TRYT_DROME	TRYPsin THETA PRECURSO	2.07e+02	323	1	CA19_RAT	COLLAGEN ALPHA 1(XI) C	2.07e+02
251	1	TRYT_DROME	TRYPsin THETA PRECURSO	2.07e+02	324	1	PERJ_ORISA	PEROXIDASE PRECURSOR (	2.07e+02
252	1	YC43_ODOSI	HYPOTHETICAL 30.1 KD P	2.07e+02	325	1	GCL_RAT	IG GAMMA-1 CHAIN C REG	2.07e+02
253	1	GP3D_CHLTR	VIRULENCE PROTEIN PG3 P	2.07e+02	326	1	GSPK_ECOLI	PROBABLE GENERAL SECRE	2.07e+02
254	1	CTRA_GADMO	CHYMOTRYPSIN A PRECURS	2.07e+02	327	1	YQ51_CAEEL	PUTATIVE MITOCHONDRIAL	2.07e+02
255	1	RM41_YEAST	MITOCHONDRIAL 60S RIBO	2.07e+02	328	1	YV22_MYCTU	HYPOTHETICAL 34.9 KD P	2.07e+02
256	1	DIVB_BACSU	DIVISION INITIATION PR	2.07e+02	329	1	GC3_MOUSE	IG GAMMA-3 CHAIN C REG	2.07e+02
257	1	MOPA_RHOCA	MOLYBDENUM-PYRIN BIND	2.07e+02	330	1	AAPC_PENCL	POSSIBLE APOSPORY-ASSO	2.07e+02
258	1	CB22_MAIZE	CHLOROPHYLL A-B BINDIN	2.07e+02	331	1	GCC_RAT	IG GAMMA-2C CHAIN C RE	2.07e+02
259	1	Y43C_ARATH	14-3-3-LIKE PROTEIN GF	2.07e+02	332	1	AP1_PIG	TRANSCRIPTION FACTOR A	2.07e+02
260	1	VLDC_SCHPO	HYPOTHETICAL 29.9 KD P	2.07e+02	333	1	VAL_BPSP	READTHROUGH PROTEIN A1	2.07e+02
261	1	COX3_PODAN	CYTOCHROME C OXIDASE P	2.07e+02	334	1	AP1_HUMAN	TRANSCRIPTION FACTOR A	2.07e+02
262	1	CXB3_HUMAN	GAP JUNCTION BETA-3 PR	2.07e+02	335	1	VS08_ROTAPC	GLYCOPROTEIN VP7 PRECU	2.07e+02
263	1	SULD_STRPN	BIFUNCTIONAL FOLATE SY	2.07e+02	336	1	VS08_ROTAPC	GLYCOPROTEIN VP7 PRECU	2.07e+02
264	1	HIS6_ARCFU	HISF PROTEIN (CYCLASE)	2.07e+02	337	1	TALI_KLULA	TRANSALDOLASE (EC 2.2.	2.07e+02
265	1	ATP6_ECOLI	ATP SYNTHASE A CHAIN (	2.07e+02	338	1	ETFA_RAT	ELECTRON TRANSFER FLAV	2.07e+02
266	1	CXB4_RAT	GAP JUNCTION BETA-4 PR	2.07e+02	339	1	YXJG_BACSU	HYPOTHETICAL 38.0 KD P	2.07e+02
267	1	NPH1_RAT	NEUREXOPHILIN 1 (NEUR	2.07e+02	340	1	AP1_RAT	TRANSCRIPTION FACTOR A	2.07e+02
268	1	ATP6_BUCAP	ATP SYNTHASE A CHAIN (	2.07e+02	341	1	AP1_MOUSE	TRANSCRIPTION FACTOR A	2.07e+02
269	1	TC1A_CAEEL	TRANSPORABLE ELEMENT T	2.07e+02	342	1	ACT4_LYCES	ACTIN 105 (FRAGMENT)	2.07e+02
270	1	MD12_SCHPO	MITOCHONDRIAL INHERITA	2.07e+02	343	1	STSP_STAAR	GLUTAMYL ENDOPEPTIDASE	2.07e+02
271	1	FOLD_MYCGE	METHYLENETHETRAHYDROFOL	2.07e+02	344	1	G3P_METTH	GLYCERALDEHYDE 3-PHOSP	2.07e+02
272	1	NIFU_KLEPN	NIFU PROTEIN	2.07e+02	345	1	G3P_CURLU	GLYCERALDEHYDE 3-PHOSP	2.07e+02
273	1	BLAC_KLEPN	BETA-LACTAMASE PRECURS	2.07e+02	346	1	G3P_PODAN	GLYCERALDEHYDE 3-PHOSP	2.07e+02
274	1	RFB5_SALTI	PARATOSE SYNTHASE (EC	2.07e+02	347	1	RI07_ACACA	MITOCHONDRIAL RIBOSOMA	2.07e+02
275	1	ATND_RAT	SODIUM/POTASSIUM-TRANS	2.07e+02	348	1	HMRO_DROVI	HOMEOBOX PROTEIN ROUGH	2.07e+02
276	1	Y005_MYCTU	HYPOTHETICAL 29.1 KD P	2.07e+02	349	1	YXJH_BACSU	HYPOTHETICAL 38.3 KD P	2.07e+02
277	1	LPXA_BRUAB	ACYL-LACYL-CARRIER-PRO	2.07e+02	350	1	OTCA_LACPL	ORNITHINE CARBAMOYLTRA	2.07e+02
278	1	BCHO_RHOCA	MAGNESIUM-CHELATASE 30	2.07e+02	351	1	UL20_HCMVA	HYPOTHETICAL PROTEIN U	2.07e+02
279	1	GSPC_KLEPN	GENERAL SECRETION PATH	2.07e+02	352	1	Y0BX_MYCTU	HYPOTHETICAL 38.0 KD P	2.07e+02
280	1	FLA3_TREPA	FLAGELLAR FILAMENT 31	2.07e+02	353	1	NKR2_HUMAN	MHC CLASS I NK CELL RE	2.07e+02
281	1	HUS1_SCHPO	HUS1 PROTEIN	2.07e+02	354	1	YM22_YEAST	HYPOTHETICAL 38.6 KD P	2.07e+02
282	1	HOG3_HORVU	GAMMA-HORDEIN 3.	2.07e+02	355	1	RECA_GLUOX	RECA PROTEIN	2.07e+02
283	1	YD53_SYNY3	HYPOTHETICAL 31.3 KD P	2.07e+02	356	1	MDH_METJA	MALATE DEHYDROGENASE (	2.07e+02
284	1	RCGR_HUMAN	RPE-RETINAL G PROTEIN-	2.07e+02	357	1	GAS1_HUMAN	GROWTH-ARREST-SPECIFIC	2.07e+02
285	1	CGD3_HUMAN	GL/S-SPECIFIC CYCLIN D	2.07e+02	358	1	Y043_BORBU	HYPOTHETICAL PROTEIN B	2.07e+02
286	1	CGD3_MOUSE	GL/S-SPECIFIC CYCLIN D	2.07e+02	359	1	PBX1_MOUSE	PRE-B-CELL LEUKEMIA TR	2.07e+02
287	1	CGD3_RAT	GL/S-SPECIFIC CYCLIN D	2.07e+02	360	1	SRD2_CAEEL	SRD-2 PROTEIN	2.07e+02
288	1	BLAC_RHOCA	BETA-LACTAMASE PRECURS	2.07e+02	361	1	RECA_ACEPO	RECA PROTEIN	2.07e+02
289	1	FENR_SPISP	FERREDOXIN--NADP REDUC	2.07e+02	362	1	NKR0_HUMAN	MHC CLASS I NK CELL RE	2.07e+02
290	1	DRAG_RHORA	ADP-RIBOSYLGLYCOPHOL	2.07e+02	363	1	YJN6_CAEEL	HYPOTHETICAL 38.0 KD P	2.07e+02
291	1	DAPF_METJA	DIAMINOPTIMELATE EPIMER	2.07e+02	364	1	Y239_METJA	HYPOTHETICAL PROTEIN M	2.07e+02
292	1	REBF_SHIFL	DTDP-RHAMNOSYL TRANSFE	2.07e+02	365	1	SMP_SERMA	EXTRACELLULAR MINOR ME	2.07e+02
293	1	YMY9_YEAST	HYPOTHETICAL 34.0 KD P	2.07e+02	366	1	YQHT_BACSU	PUTATIVE PEPTIDASE IN	2.07e+02
294	1	NIFH_NOSCO	NITROGENASE IRON PROTE	2.07e+02	367	1	T2BA_BACAR	TYPE II RESTRICTION EN	2.07e+02
295	1	LPDP_PSEAP	LIPOPROTEIN NLDP/LPPB	2.07e+02	368	1	LIMA_PSEGL	LIPASE MODULATOR PRECU	2.07e+02
296	1	AOX2_TOBAC	ALTERNATIVE OXIDASE 2	2.07e+02	369	1	CA29_CHICK	COLLAGEN ALPHA 2(IX) C	2.07e+02
297	1	MPEV_SYNDP	BILIN BIOSYNTHESIS PRO	2.07e+02	370	1	CHLI_GUITH	MAGNESIUM-CHELATASE SU	2.07e+02
298	1	NKR5_HUMAN	MHC CLASS I NK CELL RE	2.07e+02	371	1	FENR_CHDRE	FERREDOXIN--NADP REDUC	2.07e+02
299	1	CSVR_ECOLI	TRANSCRIPTIONAL ACTIVA	2.07e+02	372	1	TYRA_LACIA	PREPHEPNATE DEHYDROGEN	2.07e+02
300	1	PCYB_PSEPA	PROTEOGUATE 4,5-DI	2.07e+02	373	1	KKR3_MACMU	C-C CHEMOKINE RECEPTOR	2.07e+02
301	1	NKR9_HUMAN	MHC CLASS I NK CELL RE	2.07e+02	374	1	CKR3_CERAE	C-C CHEMOKINE RECEPTOR	2.07e+02
302	1	Y031_STRGR	HYPOTHETICAL 31.2 KD P	2.07e+02	375	1	FL3H_GALCH	NARINGENIN, 2-OXOGLUTAR	2.07e+02
303	1	FMT1_THERT	METHIONYL-TRNA FORMYLT	2.07e+02	376	1	TF2B_KLULA	TRANSCRIPTION INITIATI	2.07e+02
304	1	PJA_HELAN	SERINE/THREONINE PROTE	2.07e+02	377	1	WN8B_BRARE	WNT-8B PROTEIN PRECURS	2.07e+02
305	1	NUBM_HUMAN	NADH-UBIQUINONE OXIDOR	2.07e+02	378	1	HIS8_LACIA	HISTIDINOL-PHOSPHATE A	2.07e+02
306	1	YK53_YEAST	HYPOTHETICAL 34.3 KD P	2.07e+02	379	1	Y446_METJA	HYPOTHETICAL PROTEIN M	2.07e+02
307	1	P2AL_ARATH	SERINE/THREONINE PROTE	2.07e+02	380	1	RL4B_YEAST	60S RIBOSOMAL PROTEIN	2.07e+02
308	1	P2AL_ARATH	SERINE/THREONINE PROTE	2.07e+02	381	1	ELY2_BACSP	60S RIBOSOMAL PROTEIN	2.07e+02
309	1	SPFH_BOVIN	SYNAPTOPHYSIN (MAJOR S	2.07e+02	382	1	ASK1_ARATH	THERMOSTABLE ALKALINE	2.07e+02
310	1	ADTL1_YEAST	ADP,ATP CARRIER PROTEI	2.07e+02	383	1	CYB_TRYBTH	SERINE/THREONINE-PROTE	2.07e+02
311	1	YF56_METJA	HYPOTHETICAL PROTEIN M	2.07e+02	384	1	CYB_TRYBTH	CYCLOTHROMB B (EC 1.10.	2.07e+02
312	1	RPOA_ODOSI	DNA-DIRECTED RNA POLYM	2.07e+02	385	1	PGLR_COCCA	POLYGALACTURONASE PREC	2.07e+02
313	1	COLY_YERPE	COAGULASE/FIBRINOLYSIN	2.07e+02	386	1	FL3H_VITVI	NARINGENIN, 2-OXOGLUTAR	2.07e+02
314	1	SPFH_HUMAN	SYNAPTOPHYSIN (MAJOR S	2.07e+02	387	1	RM02_KLULA	MITOCHONDRIAL 60S RIBO	2.07e+02
315	1	RPOA_PYRSA	DNA-DIRECTED RNA POLYM	2.07e+02	388	1	DCUP_RAT	UROPOPHRYNOGEN DECAR	2.07e+02

315	1	YOH1_ECOLI	HYPOTHETICAL 35.2 KD P	2.07e+02
317	1	CAIC_RAT	COLLAGEN ALPHA 1(XII)	2.07e+02
320	1	KLAO_SOYBN	60S ACIDIC RIBOSOMAL P	2.07e+02
321	1	FCEE2_HUMAN	LOW AFFINITY IMMUNOGLO	2.07e+02
322	1	RM10_YEAST	MITOCHONDRIAL 60S RIBO	2.07e+02
322	1	GCA_RAT	IG GAMMA-2A CHAIN C RE	2.07e+02
323	1	QOX2_BACSU	QUINOL OXIDASE POLYPEP	2.07e+02
323	1	CA19_RAT	COLLAGEN ALPHA 1(XI) C	2.07e+02
325	1	PERJ_ORISA	PEROXIDASE PRECURSOR (	2.07e+02
326	1	GCL_RAT	IG GAMMA-1 CHAIN C REG	2.07e+02
327	1	GSPK_ECOLI	PROBABLE GENERAL SECRE	2.07e+02
328	1	YQ51_CAEEL	PUTATIVE MITOCHONDRIAL	2.07e+02
328	1	YV22_MYCTU	HYPOTHETICAL 34.9 KD P	2.07e+02
329	1	GC3_MOUSE	IG GAMMA-3 CHAIN C REG	2.07e+02
329	1	AAPC_PENCL	POSSIBLE APOSPORY-ASSO	2.07e+02
329	1	GCC_RAT	IG GAMMA-2C CHAIN C RE	2.07e+02
331	1	AP1_PIG	TRANSCRIPTION FACTOR A	2.07e+02
331	1	VAL_BPSP	READTHROUGH PROTEIN A1	2.07e+02
331	1	AP1_HUMAN	TRANSCRIPTION FACTOR A	2.07e+02
332	1	VS08_ROTAPC	GLYCOPROTEIN VP7 PRECU	2.07e+02
332	1	VS08_ROTAPC	GLYCOPROTEIN VP7 PRECU	2.07e+02
333	1	TALI_KLULA	TRANSALDOLASE (EC 2.2.	2.07e+02
333	1	ETFA_RAT	ELECTRON TRANSFER FLAV	2.07e+02
334	1	YXJG_BACSU	HYPOTHETICAL 38.0 KD P	2.07e+02
334	1	AP1_RAT	TRANSCRIPTION FACTOR A	2.07e+02
334	1	AP1_MOUSE	TRANSCRIPTION FACTOR A	2.07e+02
336	1	ACT4_LYCES	ACTIN 105 (FRAGMENT)	2.07e+02
336	1	STSP_STAAR	GLUTAMYL ENDOPEPTIDASE	2.07e+02
337	1	G3P_METTH	GLYCERALDEHYDE 3-PHOSP	2.07e+02
337	1	G3P_CURLU	GLYCERALDEHYDE 3-PHOSP	2.07e+02
337	1	G3P_PODAN	GLYCERALDEHYDE 3-PHOSP	2.07e+02
337	1	RI07_ACACA	MITOCHONDRIAL RIBOSOMA	2.07e+02
339	1	HMRO_DROVI	HOMEOBOX PROTEIN ROUGH	2.07e+02
340	1	YXJH_BACSU	HYPOTHETICAL 38.3 KD P	2.07e+02
340	1	OTCA_LACPL	ORNITHINE CARBAMOYLTRA	2.07e+02
340	1	UL20_HCMVA	HYPOTHETICAL PROTEIN U	2.07e+02
341	1	Y0BX_MYCTU	HYPOTHETICAL 38.0 KD P	2.07e+02
341	1	NKR2_HUMAN	MHC CLASS I NK CELL RE	2.07e+02
342	1	YM22_YEAST	HYPOTHETICAL 38.6 KD P	2.07e+02
342	1	RECA_GLUOX	RECA PROTEIN	2.07e+02
344	1	MDH_METJA	MALATE DEHYDROGENASE (	2.07e+02
345	1	GAS1_HUMAN	GROWTH-ARREST-SPECIFIC	2.07e+02
346	1	Y043_BORBU	HYPOTHETICAL PROTEIN B	2.07e+02
347	1	PBX1_MOUSE	PRE-B-CELL LEUKEMIA TR	2.07e+02
347	1	SRD2_CAEEL	SRD-2 PROTEIN	2.07e+02
348	1	RECA_ACEPO	RECA PROTEIN	2.07e+02
348	1	NKR0_HUMAN	MHC CLASS I NK CELL RE	2.07e+02
349	1	YJN6_CAEEL	HYPOTHETICAL 38.0 KD P	2.07e+02
351	1	Y239_METJA	HYPOTHETICAL PROTEIN M	2.07e+02
352	1	SMP_SERMA	EXTRACELLULAR MINOR ME	2.07e+02
353	1	YQHT_BACSU	PUTATIVE PEPTIDASE IN	2.07e+02
353	1	T2BA_BACAR	TYPE II RESTRICTION EN	2.07e+02
353	1	LIMA_PSEGL	LIPASE MODULATOR PRECU	2.07e+02
353	1	CA29_CHICK	COLLAGEN ALPHA 2(IX) C	2.07e+02
353	1	CHLI_GUITH	MAGNESIUM-CHELATASE SU	2.07e+02
353	1	FENR_CHDRE	FERREDOXIN--NADP REDUC	2.07e+02
354	1	TYRA_LACIA	PREPHEPNATE DEHYDROGEN	2.07e+02
355	1	KKR3_MACMU	C-C CHEMOKINE RECEPTOR	2.07e+02
355	1	CKR3_CERAE	C-C CHEMOKINE RECEPTOR	2.07e+02
356	1	FL3H_GALCH	NARINGENIN, 2-OXOGLUTAR	2.07e+02
357	1	TF2B_KLULA	TRANSCRIPTION INITIATI	2.07e+02
358	1	WN8B_BRARE	WNT-8B PROTEIN PRECURS	2.07e+02
360	1	HIS8_LACIA	HISTIDINOL-PHOSPHATE A	2.07e+02
361	1	Y446_METJA	HYPOTHETICAL PROTEIN M	2.07e+02
361	1	RL4B_YEAST	60S RIBOSOMAL PROTEIN	2.07e+02
361	1	ELY2_BACSP	60S RIBOSOMAL PROTEIN	2.07e+02
361	1	ASK1_ARATH	THERMOSTABLE ALKALINE	2.07e+02
363	1	CYB_TRYBTH	SERINE/THREONINE-PROTE	2.07e+02
363	1	CYB_TRYBTH	CYCLOTHROMB B (EC 1.10.	2.07e+02
364	1	PGLR_COCCA	POLYGALACTURONASE PREC	2.07e+02
364	1	FL3H_VITVI	NARINGENIN, 2-OXOGLUTAR	2.07e+02
364	1	RM02_KLULA	MITOCHONDRIAL 60S RIBO	2.07e+02
364	1	DCUP_RAT	UROPOPHRYNOGEN DECAR	2.07e+02







681	1	NOP2_YEAST	NUCLEOLAR PROTEIN NOP2	2.07e+02	754	1	RA15_SCHPO	DNA REPAIR HELICASE RA	2.07e+02
682	1	CHIT_STRLI	CHITINASE C PRECURSOR	2.07e+02	755	1	MUS2_BORBU	MUS22 PROTEIN.	2.07e+02
683	1	YM20_YEAST	HYPOTHETICAL 72.2 KD P	2.07e+02	756	1	PRTP_HSV1F	PROCESSING AND TRANSP	2.07e+02
684	1	YF54_METJA	HYPOTHETICAL PROTEIN M	2.07e+02	757	1	CAD5_PIG	VASCULAR ENDOTHELIAL-C	2.07e+02
685	1	OGP_PAPAN	OVIDUCT-SPECIFIC GLYCO	2.07e+02	758	1	OPT3_CAEEL	HYPOTHETICAL OLIGOPEPT	2.07e+02
686	1	K08G_MYCTU	PROBABLE SERINE/THREON	2.07e+02	759	1	PRTP_HSV11	PROCESSING AND TRANSP	2.07e+02
687	1	Z195_HUMAN	ZINC FINGER PROTEIN 19	2.07e+02	760	1	PRTP_HSV1A	PROCESSING AND TRANSP	2.07e+02
688	1	MSM4_YEAST	ZINC FINGER PROTEIN MS	2.07e+02	761	1	MUS2_BACSU	MUS22 PROTEIN.	2.07e+02
689	1	FIG1_MOUSE	FIG-1 PROTEIN PRECURSO	2.07e+02	762	1	PEPX_LACDL	XAA-PRO DIPEPTIDYL-PEP	2.07e+02
690	1	YA28_SCHPO	HYPOTHETICAL 71.9 KD P	2.07e+02	763	1	REGA_DICDI	3',5'-CYCLIC-NUCLEOTID	2.07e+02
691	1	PBP2_ECOLI	PENICILLIN-BINDING PRO	2.07e+02	764	1	FL08_YEAST	TRANSCRIPTIONAL ACTIVA	2.07e+02
692	1	DI11_RAT	DYNEIN INTERMEDIATE CH	2.07e+02	765	1	PT27_YEAST	PUTATIVE MITOCHONDRIAL	2.07e+02
693	1	VP40_HSVB	CAPSID PROTEIN P40 (VI	2.07e+02	766	1	ACSB_ACEXY	CELLULOSE SYNTHASE 93	2.07e+02
694	1	TKT_MYCPN	TRANSKETOLASE (EC 2.2.	2.07e+02	767	1	FPS_DROME	TYROSINE-PROTEIN KINAS	2.07e+02
695	1	TKT_MYCGE	TRANSKETOLASE (EC 2.2.	2.07e+02	768	1	SYFB_SYNP7	PHENYLALANYL-TRNA SYN	2.07e+02
696	1	PRTS_MACMU	VITAMIN K-DEPENDENT PR	2.07e+02	769	1	PUPB_PSEPU	FERRIC-PSEUDOBACTIN BN	2.07e+02
697	1	DMR9_MOUSE	DMR-N9 PROTEIN.	2.07e+02	770	1	YJF1_YEAST	HYPOTHETICAL 91.3 KD P	2.07e+02
698	1	MTS1_STRSA	MODIFICATION METHYLASE	2.07e+02	771	1	VSUB_HSV7J	GLYCOPROTEIN B PRECURS	2.07e+02
699	1	TKT_STRPN	PROBABLE TRANSKETOLASE	2.07e+02	772	1	MSU2_SYNY3	MUS22 PROTEIN.	2.07e+02
700	1	CSPI_CORGL	PSI PROTEIN PRECURSOR.	2.07e+02	773	1	CC27_HUMAN	PROTEIN CDC27HS (CELL	2.07e+02
701	1	LAMA_CHICK	LAMIN A.	2.07e+02	774	1	TRKC_PIG	NT-3 GROWTH FACTOR REC	2.07e+02
702	1	G390_MYCPN	HYPOTHETICAL ATP-BINDI	2.07e+02	775	1	PMFC_PROMI	OUTER MEMBRANE USHER P	2.07e+02
703	1	GSPD_KLEPN	GENERAL SECRETION PATH	2.07e+02	776	1	KL91_SCHPO	KINESIN-LIKE PROTEIN 1	2.07e+02
704	1	FHUB_ECOLI	FERRICHRONE TRANSPORT	2.07e+02	777	1	MSH5_HUMAN	MUS22 PROTEIN HOMOLOG 5	2.07e+02
705	1	PAB5_MOUSE	HISTIDINE DECARBOXYLAS	2.07e+02	778	1	INVA_YEREN	INVASIN.	2.07e+02
706	1	PAB5_ARATH	POLYADENYLATE-BINDING	2.07e+02	779	1	TGLK_RABIT	PROTEIN-GLUTAMINE GAMM	2.07e+02
707	1	SR72_CANFA	SIGNAL RECOGNITION PAR	2.07e+02	780	1	TRKC_HUMAN	NT-3 GROWTH FACTOR REC	2.07e+02
708	1	SR72_HUMAN	SIGNAL RECOGNITION PAR	2.07e+02	781	1	HEX_VIBU	BETA-HEXOSAMINIDASE (E	2.07e+02
709	1	CA1A_BOVIN	COLLAGEN 1(X) CHAIN PR	2.07e+02	782	1	C22B_HUMAN	B-CELL RECEPTOR CD22-B	2.07e+02
710	1	RP5D_RHOCA	RNA POLYMERASE SIGMA F	2.07e+02	783	1	PR05_YEAST	PRE-mRNA PROCESSING RN	2.07e+02
711	1	CA1A_CHICK	COLLAGEN 1(X) CHAIN PR	2.07e+02	784	1	TRKC_CHICK	NT-3 GROWTH FACTOR REC	2.07e+02
712	1	PRTS_HUMAN	VITAMIN K-DEPENDENT PR	2.07e+02	785	1	MUS2_SALTU	DNA MISMATCH REPAIR PR	2.07e+02
713	1	OGP_HUMAN	OVIDUCT-SPECIFIC GLYCO	2.07e+02	786	1	Y211_MYCTU	PROBABLE CATION-TRANSP	2.07e+02
714	1	YKR9_YEAST	HYPOTHETICAL 77.5 KD P	2.07e+02	787	1	VPH1_NEUCR	VACUOLAR ATP SYNTHASE	2.07e+02
715	1	NOLO_RHISN	MODULATION PROTEIN NOL	2.07e+02	788	1	HMDH_BLAG	3-HYDROXY-3-METHYLGLUT	2.07e+02
716	1	CA1A_HUMAN	COLLAGEN ALPHA 1(X) CH	2.07e+02	789	1	NIA_VOICA	NITRATE REDUCTASE (EC	2.07e+02
717	1	CA1A_MOUSE	COLLAGEN ALPHA 1(X) CH	2.07e+02	790	1	TRKC_RAT	NITRATE REDUCTASE (EC	2.07e+02
718	1	CY2G_RAT	GUANYLATE CYCLASE SOLU	2.07e+02	791	1	CPN_DROME	CALPHOTIN.	2.07e+02
719	1	LIPE_AERYH	EXTRACELLULAR LIPASE P	2.07e+02	792	1	P100_HSV6U	LARGE STRUCTURAL PHOSP	2.07e+02
720	1	HMC1_HELPY	PROBABLE HEAVY-METAL C	2.07e+02	793	1	DP01_SULOH	DNA POLYMERASE I (EC 2	2.07e+02
721	1	MTS_PODCA	MYOSIN HEAVY CHAIN (FR	2.07e+02	794	1	DNAB_SYNY3	REPLICATIVE DNA HELICA	2.07e+02
722	1	CAUP_DROME	HOMEBOX PROTEIN CAUPO	2.07e+02	795	1	VGLB_ILTV6	GLYCOPROTEIN B PRECURS	2.07e+02
723	1	YAB5_SCHPO	HYPOTHETICAL 81.8 KD P	2.07e+02	796	1	MGR3_HUMAN	METABOTROPIC GLUTAMATE	2.07e+02
724	1	SECD_MYCLE	PROTEIN-EXPORT MEMBRAN	2.07e+02	797	1	SECA_ANTSP	PREPROTEIN TRANSLOCASE	2.07e+02
725	1	YCBY_ECOLI	HYPOTHETICAL 78.9 KD P	2.07e+02	798	1	Y218_MYCTU	PUTATIVE REGULATORY PR	2.07e+02
726	1	HS83_YEAST	HEAT SHOCK COGNATE PRO	2.07e+02	799	1	PMI1_HUMAN	PROBABLE TRANSCRIPTION	2.07e+02
727	1	PLB2_YEAST	PUTATIVE LYOPHOSPHOLI	2.07e+02	800	1	AREA_ASPNG	NITROGEN REGULATORY PR	2.07e+02
728	1	HLXB_PROVU	HEMOLYSIN SECRETION AT	2.07e+02	801	1	VGLB_ILTVS	GLYCOPROTEIN B PRECURS	2.07e+02
729	1	HLXB_ACTAC	LEUKOTOXIN SECRETION A	2.07e+02	802	1	VGLB_ILTVT	GLYCOPROTEIN B PRECURS	2.07e+02
730	1	HLXB_PASPP	LEUKOTOXIN SECRETION A	2.07e+02	803	1	APCE_AGLNE	PHYCOBILISOME LINKER P	2.07e+02
731	1	HLXB_PASHA	LEUKOTOXIN SECRETION A	2.07e+02	804	1	ITH3_HUMAN	INTER-ALPHA-TRYPSIN IN	2.07e+02
732	1	TE21_NEIME	TRANSFERRIN-BINDING PR	2.07e+02	805	1	ITH3_MOUSE	INTER-ALPHA-TRYPSIN IN	2.07e+02
733	1	RECA_MYCLE	RECA PROTEIN.	2.07e+02	806	1	ITH3_RAT	INTER-ALPHA-TRYPSIN IN	2.07e+02
734	1	E2BE_YEAST	TRANSLATION INITIATION	2.07e+02	807	1	YA54_METJA	HYPOTHETICAL PROTEIN M	2.07e+02
735	1	ARA_DROME	HOMEBOX PROTEIN ARAUC	2.07e+02	808	1	A180_MOUSE	CLATHRIN COAT ASSEMBLY	2.07e+02
736	1	PALY_RHOTO	PHENYLALANINE AMMONIA-	2.07e+02	809	1	SYLM_HUMAN	PROBABLE LEUCYL-TRNA S	2.07e+02
737	1	YK4A_CAEEL	HYPOTHETICAL 82.8 KD P	2.07e+02	907	1	ITH1_MOUSE	INTER-ALPHA-TRYPSIN IN	2.07e+02
738	1	Y022_TREPA	HYPOTHETICAL PROTEIN T	2.07e+02	907	1	LEMA_PSEY	SENSOR PROTEIN LEMA (E	2.07e+02
739	1	YP65_CAEEL	HYPOTHETICAL 81.5 KD P	2.07e+02	914	1	ITH1_MESAU	INTER-ALPHA-TRYPSIN IN	2.07e+02
740	1	PCRA_BACST	ATP-DEPENDENT HELICASE	2.07e+02	915	1	A180_RAT	CLATHRIN COAT ASSEMBLY	2.07e+02
741	1	CTCF_HUMAN	TRANSCRIPTIONAL REPRES	2.07e+02	916	1	HMDH_DROME	3-HYDROXY-3-METHYLGLUT	2.07e+02
742	1	ACET_HUMAN	ANGIOTENSIN-CONVERTING	2.07e+02	920	1	NIA_CICIN	NITRATE REDUCTASE (EC	2.07e+02
743	1	YH09_YEAST	HYPOTHETICAL 81.8 KD P	2.07e+02	930	1	ITH4_HUMAN	INTER-ALPHA-TRYPSIN IN	2.07e+02
744	1	PSAB_CHLUV	PHOTOSYSTEM I P700 CHL	2.07e+02	931	1	CA19_HUMAN	COLLAGEN ALPHA 1(IX) C	2.07e+02
745	1	PSAB_PORPU	PHOTOSYSTEM I P700 CHL	2.07e+02	932	1	PM51_HUMAN	PMS1 PROTEIN HOMOLOG 1	2.07e+02
746	1	YN02_CAEEL	HYPOTHETICAL 84.0 KD P	2.07e+02	935	1	ITH2_PIG	INTER-ALPHA-TRYPSIN IN	2.07e+02
747	1	ATI12_HSVB	ALPHA TRANS-INDUCING F	2.07e+02	943	1	YLW5_CAEEL	HYPOTHETICAL 105.9 KD	2.07e+02
748	1	FHUA_ECOLI	FERRICHRONE-IRON RECEP	2.07e+02	943	1	IROA_NEIME	IRON-REGULATED OUTER M	2.07e+02
749	1	Y226_HUMAN	HYPOTHETICAL PROTEIN K	2.07e+02	947	1	HLYA_PASPP	LEUKOTOXIN.	2.07e+02
750	1	NEC1_RAT	NEUROENDOCRINE CONVERT	2.07e+02	947	1	BGLS_RUMAL	BETA-GLUCOSIDASE (EC 3	2.07e+02
751	1	CLPA_ECOLI	ATP-DEPENDENT CLP PROT	2.07e+02	951	1	HEX_ADE05	HEXON PROTEIN (LATE PR	2.07e+02
752	1	YKH2_YEAST	HYPOTHETICAL 88.8 KD P	2.07e+02	952	1	YK15_CAEEL	HYPOTHETICAL 105.3 KD	2.07e+02
753	1	SWI6_KLULA	REGULATORY PROTEIN SWI	2.07e+02	956	1	GCSP_ECOLI	GLYCINE DEHYDROGENASE	2.07e+02



827	1	0.7	956	1	0.7	900	1	1332	1	SPT7_YEAST	TRANSCRIPTIONAL ACTIVA	2.07e+02
828	1	0.7	958	1	0.7	901	1	1337	1	DEX1_STRDO	DEXTRANASE PRECURSOR	2.07e+02
829	1	0.7	967	1	0.7	902	1	1383	1	INSR_RAT	INSULIN RECEPTOR PREC	2.07e+02
830	1	0.7	969	1	0.7	903	1	1388	1	CALE_HUMAN	COLLAGEN ALPHA 1(XV) C	2.07e+02
831	1	0.7	970	1	0.7	904	1	1389	1	PRAX_RAT	PERIAxin	2.07e+02
832	1	0.7	971	1	0.7	905	1	1409	1	RPOC_PSEPU	DNA-DIRECTED RNA POLYM	2.07e+02
833	1	0.7	976	1	0.7	906	1	1411	1	TCOF_HUMAN	TREACLE PROTEIN (TREAC	2.07e+02
834	1	0.7	976	1	0.7	907	1	1420	1	YMBE_YEAST	HYPOTHETICAL 163.6 KD	2.07e+02
835	1	0.7	985	1	0.7	908	1	1426	1	TOP_DROME	GURKEN RECEPTOR PREC	2.07e+02
836	1	0.7	987	1	0.7	909	1	1427	1	REST_HUMAN	RESTIN (CYTOPLASMIC LI	2.07e+02
837	1	0.7	991	1	0.7	910	1	1430	1	GTDF_STRMO	GLUCOSYLTRANSFERASE-S	2.07e+02
838	1	0.7	992	1	0.7	911	1	1444	1	ADPL_MYCGE	ADHESIN P1 PRECURSOR	2.07e+02
839	1	0.7	992	1	0.7	912	1	1450	1	MPSE_CHICK	M-PROTEIN, STRATED MU	2.07e+02
840	1	0.7	994	1	0.7	913	1	1451	1	DPOA_RAT	DNA POLYMERASE ALPHA C	2.07e+02
841	1	0.7	999	1	0.7	914	1	1451	1	VGL2_CVCAI	E2 GLYCOPROTEIN PREC	2.07e+02
842	1	0.7	1004	1	0.7	915	1	1468	1	DPOA_YEAST	DNA POLYMERASE ALPHA C	2.07e+02
843	1	0.7	1005	1	0.7	916	1	1475	1	APU_THETI	AMYLPOULLULANASE PREC	2.07e+02
844	1	0.7	1006	1	0.7	917	1	1475	1	TRA2_CAEEL	SEX-DETERMINING TRANS	2.07e+02
845	1	0.7	1006	1	0.7	918	1	1481	1	APU_THETI	AMYLPOULLULANASE PREC	2.07e+02
846	1	0.7	1018	1	0.7	919	1	1513	1	STUI_YEAST	MITOTIC SPINDLE PROTEI	2.07e+02
847	1	0.7	1029	1	0.7	920	1	1522	1	YB6A_SCHPO	HYPOTHETICAL 171.5 KD	2.07e+02
848	1	0.7	1033	1	0.7	921	1	1533	1	PUM_DROME	MATERNAL PUMILIO PROTE	2.07e+02
849	1	0.7	1036	1	0.7	922	1	1541	1	IGAL_HAEIN	IMMUNOGLOBULIN A1 PROT	2.07e+02
850	1	0.7	1037	1	0.7	923	1	1549	1	TRHY_SHEEP	TRICHOVALIN	2.07e+02
851	1	0.7	1039	1	0.7	924	1	1551	1	VGLM_DUGBY	M POLYPROTEIN PRECURSO	2.07e+02
852	1	0.7	1040	1	0.7	925	1	1551	1	GLTS_SYNY3	FERREDOXIN-DEPENDENT G	2.07e+02
853	1	0.7	1041	1	0.7	926	1	1569	1	YPJA_ECOLI	HYPOTHETICAL 98.4 KD P	2.07e+02
854	1	0.7	1052	1	0.7	927	1	1569	1	GLI3_XENLA	ZINC FINGER PROTEIN GL	2.07e+02
855	1	0.7	1055	1	0.7	928	1	1570	1	HLVA_PROMI	HEMOLYSIN PRECURSOR	2.07e+02
856	1	0.7	1058	1	0.7	929	1	1587	1	SUR_RAT	SULFONYLUREA RECEPTOR	2.07e+02
857	1	0.7	1059	1	0.7	930	1	1591	1	SUR_CRICR	SULFONYLUREA RECEPTOR	2.07e+02
858	1	0.7	1064	1	0.7	931	1	1597	1	GTFT_STRDO	GLUCOSYLTRANSFERASE-I	2.07e+02
859	1	0.7	1084	1	0.7	932	1	1608	1	HLVA_SERMA	HEMOLYSIN PRECURSOR	2.07e+02
860	1	0.7	1091	1	0.7	933	1	1638	1	BRM_DROME	HOMEOTIC GENE REGULATO	2.07e+02
861	1	0.7	1102	1	0.7	934	1	1645	1	168K_RICTY	168 KD SURFACE-LAYER P	2.07e+02
862	1	0.7	1106	1	0.7	935	1	1647	1	SN24_HUMAN	POSSIBLE GLOBAL TRANS	2.07e+02
863	1	0.7	1107	1	0.7	936	1	1659	1	VIT_ONCMY	VITELLOGENIN PRECURSOR	2.07e+02
864	1	0.7	1111	1	0.7	937	1	1699	1	CA14_MOUSE	PROCOLLAGEN ALPHA 1(IV	2.07e+02
865	1	0.7	1113	1	0.7	938	1	1669	1	CA14_HUMAN	PROCOLLAGEN ALPHA 1(IV	2.07e+02
866	1	0.7	1114	1	0.7	939	1	1681	1	CLH_CAEEL	PROBABLE CLATHRIN HEAV	2.07e+02
867	1	0.7	1130	1	0.7	940	1	1685	1	CA54_HUMAN	COLLAGEN ALPHA 5(IV) C	2.07e+02
868	1	0.7	1131	1	0.7	941	1	1716	1	RPAL_RAT	DNA-DIRECTED RNA POLYM	2.07e+02
869	1	0.7	1137	1	0.7	942	1	1717	1	RPAL_MOUSE	DNA-DIRECTED RNA POLYM	2.07e+02
870	1	0.7	1141	1	0.7	943	1	1774	1	MSAS_PENPA	6-METHYLSALICYLIC ACID	2.07e+02
871	1	0.7	1146	1	0.7	944	1	1775	1	CA14_DROME	TENASCIN PRECURSOR (TN	2.07e+02
872	1	0.7	1152	1	0.7	945	1	1808	1	CUT1_SCHPO	CUT1 PROTEIN	2.07e+02
873	1	0.7	1164	1	0.7	946	1	1828	1	CBPA_CLOCL	CELLULOSE BINDING PROT	2.07e+02
874	1	0.7	1164	1	0.7	947	1	1848	1	FAS2_PENPA	FATTY ACID SYNTHASE, S	2.07e+02
875	1	0.7	1164	1	0.7	948	1	1857	1	ANK1_HUMAN	ANKYRIN R (ANKYRINS 2	2.07e+02
876	1	0.7	1168	1	0.7	949	1	1880	1	VIT1_CHICK	VITELLOGENIN I PRECURS	2.07e+02
877	1	0.7	1168	1	0.7	950	1	1912	1	MYSB_RAT	MYOSIN HEAVY CHAIN, CA	2.07e+02
878	1	0.7	1172	1	0.7	951	1	1935	1	MYSA_RAT	MYOSIN HEAVY CHAIN, CA	2.07e+02
879	1	0.7	1174	1	0.7	952	1	1938	1	MYSA_MOUSE	MYOSIN HEAVY CHAIN, CA	2.07e+02
880	1	0.7	1180	1	0.7	953	1	1938	1	MYSA_MOUSE	MYOSIN HEAVY CHAIN, CA	2.07e+02
881	1	0.7	1189	1	0.7	954	1	1939	1	MYSA_HUMAN	MYOSIN HEAVY CHAIN, CA	2.07e+02
882	1	0.7	1198	1	0.7	955	1	1940	1	MYSE_HUMAN	MYOSIN HEAVY CHAIN, FA	2.07e+02
883	1	0.7	1201	1	0.7	956	1	1940	1	MYSE_RAT	MYOSIN HEAVY CHAIN, FA	2.07e+02
884	1	0.7	1202	1	0.7	957	1	1967	1	RRPO_PVMR	RNA REPLICATION PROTEI	2.07e+02
885	1	0.7	1205	1	0.7	958	1	2067	1	BIMB_EMENI	CELL DIVISION-ASSOCIAT	2.07e+02
886	1	0.7	1216	1	0.7	959	1	2090	1	HFC1_MESAU	HOST CELL FACTOR C1 (H	2.07e+02
887	1	0.7	1239	1	0.7	960	1	2090	1	RRPL_TOSV	RNA POLYMERASE (EC 2.7	2.07e+02
888	1	0.7	1239	1	0.7	961	1	2116	1	RRPL_SYNY	BETA SU	2.07e+02
889	1	0.7	1242	1	0.7	962	1	2167	1	BM2_YEAST	GTPASE ACTIVATING PROT	2.07e+02
890	1	0.7	1250	1	0.7	963	1	2205	1	POLN_RUBVT	NONSTRUCTURAL POLYPROT	2.07e+02
891	1	0.7	1252	1	0.7	964	1	2261	1	RRPL_MUMPM	RNA POLYMERASE BETA SU	2.07e+02
892	1	0.7	1265	1	0.7	965	1	2316	1	PTPZ_RAT	PROTEIN-TYROSINE PHOSP	2.07e+02
893	1	0.7	1266	1	0.7	966	1	2333	1	POLG_FMDV1	GENOME POLYPROTEIN [CO	2.07e+02
894	1	0.7	1295	1	0.7	967	1	2344	1	POLN_RHVD	NON-STRUCTURAL POLYPRO	2.07e+02
895	1	0.7	1296	1	0.7	968	1	2405	1	DYHA_CHLRE	DYNEIN ALPHA CHAIN, FL	2.07e+02
896	1	0.7	1305	1	0.7	969	1	2703	1	NOTC_DROME	NEUROGENIC LOCUS NOTCH	2.07e+02
897	1	0.7	1306	1	0.7	970	1	2768	1	THYG_MOUSE	THYOGLOBULIN PRECURSOR	2.07e+02
898	1	0.7	1312	1	0.7	971	1	2871	1	FBN1_HUMAN	FIBRILLIN 1 PRECURSOR	2.07e+02
899	1	0.7	1323	1	0.7	972	1		1			

```
973 6 0.7 2871 1 FBW1_BOVIN 2.07e+02
974 6 0.7 2875 1 RPL_TSWV1 2.07e+02
975 6 0.7 2944 1 CAL1_HUMAN 2.07e+02
976 6 0.7 3010 1 POLG_HCVTV 2.07e+02
977 6 0.7 3010 1 POLG_HCVBK 2.07e+02
978 6 0.7 3010 1 POLG_HCVJT 2.07e+02
979 6 0.7 3011 1 POLG_HCVH 2.07e+02
980 6 0.7 3011 1 POLG_HCVI 2.07e+02
981 6 0.7 3063 1 CA1C_HUMAN 2.07e+02
982 6 0.7 3067 1 CA1C_MOUSE 2.07e+02
983 6 0.7 3110 1 LMA2_HUMAN 2.07e+02
984 6 0.7 3124 1 CA1C_CHICK 2.07e+02
985 6 0.7 3133 1 HMCT_BOMMO 2.07e+02
986 6 0.7 3163 1 GENOME_POLYPROTEIN (H 2.07e+02
987 6 0.7 3164 1 POLG_TUMVJ 2.07e+02
988 6 0.7 3176 1 CA36_HUMAN 2.07e+02
989 6 0.7 3390 1 POLG_DEN3 2.07e+02
990 6 0.7 3649 1 ACVS_NOCLA 2.07e+02
991 6 0.7 3672 1 LML2_CABEL 2.07e+02
992 6 0.7 3744 1 YHP9_YEAST 2.07e+02
993 6 0.7 3866 1 HRX_MOUSE 2.07e+02
994 6 0.7 4092 1 DYHC_YEAST 2.07e+02
995 6 0.7 4303 1 PKD1_HUMAN 2.07e+02
996 6 0.7 4393 1 PGBM_HUMAN 2.07e+02
997 6 0.7 4488 1 RRP4_CVMJH 2.07e+02
998 6 0.7 4544 1 LRP1_HUMAN 2.07e+02
999 6 0.7 4655 1 LRP2_HUMAN 2.07e+02
1000 6 0.7 5147 1 FAT_DROME 2.07e+02
```

## ALIGNMENTS

```
RESULT 1
ID ECLC_BOVIN STANDARD; PRT; 903 AA.
AC P54281;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE
DE CHANNEL).
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TRACHEA;
RX MEDLINE; 96125078.
RA CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,
RA ARRATTE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;
RT "Cloning of an epithelial chloride channel from bovine trachea.";
RL J. BIOL. CHEM. 270:31016-31026(1995).
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSCELLULAR TRANSPORT.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: TRACHEA.
CC -!- PTM: PHOSPHORYLATED BY CAM-KINASE II.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT TRANSMEM 883 903 POTENTIAL.
FT CARBOHYD 75 75 POTENTIAL.
FT CARBOHYD 278 278 POTENTIAL.
FT CARBOHYD 360 360 POTENTIAL.
FT CARBOHYD 372 372 POTENTIAL.
FT CARBOHYD 504 504 POTENTIAL.
FT CARBOHYD 515 515 POTENTIAL.
FT CARBOHYD 688 688 POTENTIAL.
FT CARBOHYD 811 811 POTENTIAL.
FT CARBOHYD 816 816 POTENTIAL.
FT CARBOHYD 842 842 POTENTIAL.
FT CARBOHYD 857 857 POTENTIAL.
SQ SEQUENCE 903 AA; 100305 MW; 6566EC1C CRC32;

Query Match 1.5%; Score 14; DB 1; Length 903;
Best Local Similarity 100.0%; Pred. No. 6.75e-20;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 150 GRAFVHEWAHLRWG 163
Qy 151 GRAFVHEWAHLRWG 164
|||||
|||||

RESULT 2
ID VG14_BPT4 STANDARD; PRT; 256 AA.
AC P11111;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)
DE NECK PROTEIN GPI4.
GN 14.
OS BACTERIOPHAGE T4.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE;
OC T4-LIKE PHAGES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE; 89263795.
RA SELIVANOV N.A., PRILIPOV A.G., MESYANZHINOV V.V.;
RT "Nucleotide sequences of bacteriophage T4 genes 13, 14 and 15.";
RL NUCLEIC ACIDS RES. 17:3583-3583(1989).
CC -!- FUNCTION: WHEN ADDED TO COMPLETED HEADS, GENE 13 AND 14
CC STRUCTURAL PROTEINS ACTIVATE THEM FOR TAIL JOINING IN VITRO
CC COMPLEMENTATION MIXTURES. THEY ARE ALSO COMPONENTS OF THE
CC NECKED TAIL.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE PROBABLE AMINO-ACID ABC TRANSPORTER PERIPLASMIC BINDING PROTEIN YATE  
 DE PRECURSOR.  
 GN YATE.  
 OS RHIZOBIUM SP. (STRAIN NGR234).  
 OG PLASMID SYM PNR234A.  
 OC BACTERIA: PROTEOBACTERIA; ALPHA SUBDIVISION; RHIZOBIACEAE GROUP;  
 OC RHIZOBIACEAE; RHIZOBIUM.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 97305956.  
 RA FREIBERG C.A., FELLAY R., BAIRROCH A., BROUGHTON W.J., ROSENTHAL A.,  
 RA PERRET X.;  
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
 RL NATURE 387:394-401(1997).  
 CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT  
 CC -1- SYSTEM Y4TEFGH FOR AN AMINO ACID.  
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING  
 CC PROTEIN FAMILY 3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000098; G2182635; -  
 DR PROSITE: PS01039; SBP\_BACTERIAL\_3; FALSE\_NEG.  
 DR PRAM: PF00497; SBP\_bac\_3; 1.  
 KW HYPOTHETICAL PROTEIN; TRANSPORT; AMINO-ACID TRANSPORT; PERIPLASMIC;  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 300 PROBABLE AMINO-ACID ABC TRANSPORTER  
 FT PERIPLASMIC BINDING PROTEIN YATE.  
 FT SEQUENCE 300 AA; 32216 MW; FC8C1F87 CRC32;  
 SQ  
 Query Match 0.9%; Score 8; DB 1; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 1.73e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 254 DLRKFN 261  
 QY 798 DLRKFN 805  
 RESULT 4  
 ID PUTA\_SALTY STANDARD; PRT; 1320 AA.  
 AC P10503;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE PROLINE DEHYDROGENASE (EC 1.5.99.8) (PROLINE OXIDASE) / DELTA-1-  
 DE PYRROLINE-5-CARBOXYLATE DEHYDROGENASE (EC 1.5.1.12) (P5C  
 DE DEHYDROGENASE).  
 GN PUTA.  
 OS SALMONELLA TYPHIMURIUM.  
 OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;  
 OC SALMONELLA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 93241961.  
 RA ALLEN S.W., SENTI-WILLIS A.E., MALOY S.R.;  
 RT "DNA sequence of the puta gene from Salmonella typhimurium: a  
 RT bifunctional membrane-associated dehydrogenase that binds DNA.";  
 RL NUCLEIC ACIDS RES. 21:1676-1676(1993).  
 RN [2]  
 RP REVISIONS.  
 RA MALOY S.R.;  
 RL SUBMITTED (APR-1994) TO EMBL/GENEBANK/DBJ DATA BANKS.  
 RN [3]

RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE: 89127131.  
 RA HAHN D.R., MYERS R.S., KENT C.R., MALOY S.R.;  
 RT "Regulation of proline utilization in Salmonella typhimurium:  
 RT molecular characterization of the put operon, and DNA sequence of the  
 RT put control region.";  
 RL MOL. GEN. GENET. 213:125-133(1988).  
 CC -1- FUNCTION: OXIDIZES PROLINE TO GLUTAMATE FOR USE AS A CARBON AND  
 CC NITROGEN SOURCE AND ALSO FUNCTION AS A TRANSCRIPTIONAL REPRESSOR  
 CC OF THE PUT OPERON.  
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + ACCEPTOR + H(2)O = (S)-1-  
 CC PYRROLINE-5-CARBOXYLATE + REDUCED ACCEPTOR.  
 CC -1- CATALYTIC ACTIVITY: 1-PYRROLINE-5-CARBOXYLATE + NAD(+) + H(2)O =  
 CC L-GLUTAMATE + NADH.  
 CC -1- COFACTOR: FAD FLAVOPROTEIN.  
 CC -1- PATHWAY: PROLINE UTILIZATION.  
 CC -1- INDUCTION: BY PROLINE, AUTOREPRESSION AND CATABOLITE REPRESSION,  
 CC AND IS POTENTIALLY NITROGEN CONTROLLED.  
 CC -1- SIMILARITY: IN THE N-TERMINAL, TO PROLINE DEHYDROGENASES.  
 CC -1- SIMILARITY: IN THE C-TERMINAL, TO ALDEHYDE DEHYDROGENASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X70843; G470180; -  
 DR EMBL: X12569; G47852; ALT\_SEQ.  
 DR PIR: S03817; S03817.  
 DR PIR: S34910; S31910.  
 DR PIR: S33716; S33716.  
 DR STYGENE; SG10321; PUTA.  
 DR PROSITE: PS00070; ALDEHYDE DEHYDR\_CYS; 1.  
 DR PROSITE: PS00687; ALDEHYDE\_DEHYDR\_GLU; 1.  
 DR PFAM: PF00171; aldedh; 1.  
 KW MULTIFUNCTIONAL ENZYME; OXIDOREDUCTASE; FLAVOPROTEIN; FAD; NAD;  
 KW TRANSCRIPTION REGULATION; REPRESSOR; DNA-BINDING; PROLINE METABOLISM.  
 FT ACT\_SITE 883 BY SIMILARITY.  
 FT ACT\_SITE 917 917 BY SIMILARITY.  
 FT SEQUENCE 1320 AA; 144188 MW; 49BFD88B CRC32;  
 SQ  
 Query Match 0.9%; Score 8; DB 1; Length 1320;  
 Best Local Similarity 100.0%; Pred. No. 1.73e-02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1228 LAKRLPAA 1235  
 QY 371 LAKRLPAA 378  
 RESULT 5  
 ID YXEE\_BACSU STANDARD; PRT; 121 AA.  
 AC P54944;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 14.7 KD PROTEIN IN IDH-DEOR INTERGENIC REGION.  
 GN YXEE OR HS74ER.  
 OS BACILLUS SUBTILIS.  
 OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
 OC BACILLUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / BGSC141;  
 RX MEDLINE: 97021444.  
 RA YOSHIDA K.-I., FUJIMURA M., YANAI N., FUJITA Y.;  
 RT "Cloning and sequencing of a 23-kb region of the Bacillus subtilis  
 RT genome between the iol and hut operons";  
 RL DNA RES. 2:295-301(1995).  
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: D45912; G1408490; -  
 DR EMBL: Z99124; E1184683; -  
 DR SUBTILIST: BG11881; YXEE.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 121 AA; 14714 MW; 390BDC5A CRC32;

Query Match 0.8%; Score 7; DB 1; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 RLNLNQ 101  
 |||||  
 QY 322 RLNLNQ 328

RESULT 6  
 ID RPB6\_ASPM2 STANDARD; PRT; 139 AA.

AC Q08399;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT 6 HOMOLOG (EC 2.7.7.6).  
 GN PMWK-CL.

OS AFRICAN SWINE FEVER VIRUS (ISOLATE MALAWI LIL 20/1) (ASFV).  
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE;  
 OC AFRICAN SWINE FEVER-LIKE VIRUSES.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93324382.  
 RA LU Z., KUTISH G.F., SUSSMAN M.D., ROCK D.L.;  
 RT "An African swine fever virus gene with a similarity to eukaryotic  
 RT RNA polymerase subunit 6".  
 RL NUCLEIC ACIDS RES. 21:2940-2940(1993).

CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.

CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N)  
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOK / EUKARYOTIC RPB6  
 CC RNA POLYMERASE SUBUNIT FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: L15411; G304101; -  
 DR PROSITE: PS01111; RNA\_POL\_K\_14KD; 1.  
 DR PFAM: PF01192; RNA\_POL\_K; 1.  
 KW DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION.  
 SQ SEQUENCE 139 AA; 15713 MW; 11963948 CRC32;

Query Match 0.8%; Score 7; DB 1; Length 139;  
 Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 55 QASSQTL 61  
 |||||  
 QY 569 QASSQTL 575

RESULT 7

ID RPB6\_ASFB7 STANDARD; PRT; 147 AA.  
 AC P42484;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT 6 HOMOLOG (EC 2.7.7.6).  
 GN C147L.  
 OS AFRICAN SWINE FEVER VIRUS (STRAIN BA7IV) (ASFV).  
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE;  
 OC AFRICAN SWINE FEVER-LIKE VIRUSES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA YANEZ R.J., RODRIGUEZ J.M., NOGAL M.L., YUSTE L., ENRIQUEZ C.,  
 RA RODRIGUEZ J.F., VINUELA E.;  
 RT "Analysis of the complete nucleotide sequence of African swine fever  
 RT virus".  
 RL VIROLOGY 208:249-278(1995).  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N)  
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPOK / EUKARYOTIC RPB6  
 CC RNA POLYMERASE SUBUNIT FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: U18466; G780439; -  
 DR PROSITE: PS01111; RNA\_POL\_K\_14KD; 1.  
 DR PFAM: PF01192; RNA\_POL\_K; 1.  
 KW DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION.  
 SQ SEQUENCE 147 AA; 16691 MW; 4EBEE3E7 CRC32;

Query Match 0.8%; Score 7; DB 1; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 51 QASSQTL 57  
 |||||  
 QY 569 QASSQTL 575

RESULT 8

ID MAL\_RAT STANDARD; PRT; 153 AA.  
 AC Q64349;  
 DT 15-JUL-1998 (REL. 36, CREATED)  
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN (17 KD MYELIN VESICULAR  
 DE PROTEIN) (MVPI7) (NS 3).  
 GN MAL.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;  
 RX MEDLINE; 96129665  
 RA KIM T., FIEDLER K., MADISON D.L., KRUEGER W.H., PFEIFFER S.E.;  
 RT "Cloning and characterization of MVPI7: a developmentally regulated  
 RT myelin protein in oligodendrocytes".  
 RL J. NEUROSCI. RES. 42:413-422(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LEWIS; TISSUE-SPINAL CORD;  
 RX MEDLINE; 95370938.  
 RA SCHAAREN-WIEMERS N., VALENZUELA D.M., FRANK M., SCHWAB M.E.;

RT "Characterization of a rat gene, rMAL, encoding a protein with four  
hydrophobic domains in central and peripheral myelin.";  
RL J. NEUROSCI. 15:5753-5764(1995).  
CC -!- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING  
CC CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.  
CC COULD BE INVOLVED IN MYELIN BIOGENESIS AND/OR MYELIN FUNCTION.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO BRAIN AND KIDNEY.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSION IS UP-REGULATED IN  
CC OLIGODENDROCYTES AND BRAIN DURING THE PERIOD OF ACTIVE  
CC MYELINATION.  
CC -!- PTM: LIPOPROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE MAL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U31367; G914968; -.  
DR EMBL; X82557; G854417; -.  
KW TRANSMEMBRANE; LIPOPROTEIN; MYELIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16758 MW; B42E4639 CRC32;  
Query Match 0.8%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Db 3 PAAASGG 9  
QY 376 PAAASGG 382  
-----  
RESULT 9  
ID MAL CANFA STANDARD; PRT; 153 AA.  
AC Q28296.  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN (VIP17 PROTEOLIPID).  
GN MAL  
OS CANIS FAMILIARIS (DOG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC CARNIVORA; FISSIPEDIA; CANIDAE; CANIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY.  
RX MEDLINE; 96140569.  
RA ZACCHETTI D., PERAENEN J., MURATA M., FIEDLER K., SIMONS K.;  
"VIP17/MAL, a proteolipid in apical transport vesicles.";  
RL FEBS LETT. 377:465-469(1995).  
CC -!- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING  
CC CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- PTM: LIPOPROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE MAL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X92505; G1054713; -.

KW TRANSMEMBRANE; LIPOPROTEIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16624 MW; 9308C3DB CRC32;  
Query Match 0.8%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Db 3 PAAASGG 9  
QY 376 PAAASGG 382  
-----  
RESULT 10  
ID MAL MOUSE STANDARD; PRT; 153 AA.  
AC Q09198; Q08819; Q09109;  
DT 15-JUL-1998 (REL. 36, CREATED)  
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN.  
GN MAL.  
OS MUS MUSCULUS (MOUSE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C; TISSUE-BRAIN;  
RX MEDLINE; 97311420.  
RA MAGYAR J.P., EBENSPERGER C., SCHAEFEN-WILMERS N., SUTER U.;  
"Myelin and lymphocyte protein (MAL/MVPI7/VIP17) and plasmolipin are  
RT members of an extended gene family.";  
RL GENE 189:269-275(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DBA/2;  
RA ALONSO M.A.;  
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- FUNCTION: COULD BE AN IMPORTANT COMPONENT IN VESICULAR TRAFFICKING  
CC CYCLING BETWEEN THE GOLGI COMPLEX AND THE APICAL PLASMA MEMBRANE.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- PTM: LIPOPROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE MAL FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Y07627; E284031; -.  
DR EMBL; Y07628; E284031; JOINED.  
DR EMBL; Y07629; E284031; JOINED.  
DR EMBL; Y07630; E284031; JOINED.  
DR EMBL; Y07626; E274594; -.  
DR EMBL; Y07812; E283229; -.  
DR MGI; MGI:892970; MAL.  
KW TRANSMEMBRANE; LIPOPROTEIN.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 93 113 POTENTIAL.  
FT TRANSMEM 128 148 POTENTIAL.  
SQ SEQUENCE 153 AA; 16595 MW; 72323D02 CRC32;  
Query Match 0.8%; Score 7; DB 1; Length 153;  
Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
-----  
Db 3 PAAASGG 9

```

QY 376 PAAASGG 382
|||||
RESULT 11
ID YR35_MYCTU STANDARD; PRT; 154 AA.
AC P71686;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 16.4 KD PROTEIN CY21B4.35.
GN MTCY21B4.35.
OS MYCOBACTERIUM TUBERCULOSIS.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIADAE;
OC ACTINOMYCETALES; CORYNEBACTERIACEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 280108; E265578; -.
DR HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 51 71 POTENTIAL.
SQ SEQUENCE 154 AA; 16383 MW; A957D449 CRC32;

Query Match 0.8%; Score 7; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.80e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 124 AIOAVDK 130
QY 843 AIOAVDK 849
|||||

RESULT 12
ID BCSD_ACEXY STANDARD; PRT; 156 AA.
AC P37719;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DE 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE CELLULOSE SYNTHASE OPERON D PROTEIN.
GN BCSD.
OS ACETOBACTER XYLINUM (ACETOBACTER PASTEURIANUS).
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; ACETOBACTERACEAE;
OC ACETOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 53582;
RA BROWN R.M.;
RL SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54676; G455537; -.
DR CELLULOSE BIOSYNTHESIS.
KW CELLULOSE BIOSYNTHESIS.
SQ SEQUENCE 156 AA; 17375 MW; 6D53E00A CRC32;

Query Match 0.8%; Score 7; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.80e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 124 AIOAVDK 130
QY 843 AIOAVDK 849
|||||

RESULT 13
ID PRIA_LYCES STANDARD; PRT; 175 AA.
AC Q08697;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PATHOGENESIS-RELATED PROTEIN 1A1 PRECURSOR (PR-1A1).
OS LYOPERSICON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; SOLANACEAE; SOLANALES; SOLANACEAE; SOLANUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. RUTGERS; TISSUE-LEAF;
RX MEDLINE; 94247356.
RA TORNERO P., CONEJERO V., VERA P.;
RT "A gene encoding a novel isoform of the PR-1 protein family from
RT tomato is induced upon viroid infection.";
RL MOL. GEN. GENET. 243:47-53(1994).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE DEFENCE REACTION OF PLANTS
CC AGAINST PATHOGENS (BY SIMILARITY).
CC -!- INDUCTION: SYNTHESIZED DURING PATHOGEN INFECTION OR OTHER STRESS-
CC RELATED RESPONSES.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X71592; G296912; -.
DR PROSITE; PS01009; SCP_AG5_PRL_SC7_1; 1.
DR PROSITE; PS01010; SCP_AG5_PRL_SC7_2; 1.
DR PFAM; PF00188; SCP; 1.
DR HSSP; P04284; 1CPE.
KW PATHOGENESIS-RELATED PROTEIN; SIGNAL; MULTIGENE FAMILY.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 175 PATHOGENESIS-RELATED PROTEIN 1A1.
FT DISULFID 65 135 BY SIMILARITY.
FT DISULFID 108 114 BY SIMILARITY.
FT DISULFID 130 144 BY SIMILARITY.
SQ SEQUENCE 175 AA; 19660 MW; 10981265 CRC32;

Query Match 0.8%; Score 7; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.80e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 33 NAARRRV 39
QY 682 NAARRRV 688
|||||

RESULT 14
ID E320_ADE03 STANDARD; PRT; 179 AA.
AC P11321;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE EARLY E3 20.1 KD GLYCOPROTEIN.
```

OS HUMAN ADENOVIRUS TYPE 3.  
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; ADENOVIRIDAE; MASTADENOVIRUS.  
RN [1]  
RA SEQUENCE FROM N.A.  
RX MEDLINE: 87219876.  
RA SIGNAS C., AKUSJARVI G., PETTERSSON U.:  
RT "Region E3 of human adenoviruses: differences between the oncogenic  
RT adenovirus-3 and the non-oncogenic adenovirus-2";  
RL GENE 50:173-184(1986).  
CC -1- FUNCTION: E3 PROTEINS SEEM TO BE DISPENSABLE FOR VIRUS GROWTH  
CC IN TISSUE CULTURE CELLS. THEY ARE POTENTIALLY IMPORTANT FOR  
CC VIRUS GROWTH UNDER SPECIAL CONDITIONS; E3 REGION MAY HELP  
CC ADENOVIRUSES TO EVADE THE IMMUNE SURVEILLANCE OF THE HOST.  
CC -1- SIMILARITY: TO ADENOVIRUS 3 E3 20.5 KD PROTEIN.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M15952; G209905; -  
DR PIR: E29500; ERAD34.  
KW EARLY PROTEIN; GLYCOPROTEIN.  
FT CARBOHYD 29 29 POTENTIAL.  
FT CARBOHYD 57 57 POTENTIAL.  
FT CARBOHYD 70 70 POTENTIAL.  
FT CARBOHYD 75 75 POTENTIAL.  
FT CARBOHYD 123 123 POTENTIAL.  
SQ SEQUENCE 179 AA; 20059 MW; 8F9A2045 CRC32;  
  
Query Match 0.8%; Score 7; DB 1; Length 179;  
Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 2 ASVTALI 8  
QY 624 ASVTALI 630  
  
RESULT 15  
ID YOY8\_CAEEL STANDARD; PRT; 194 AA.  
AC P34665;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 23.0 KD PROTEIN ZK652.8 IN CHROMOSOME III.  
GN ZK652.8.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;  
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE: 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FRASER A.,  
RA FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,  
RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,  
RA LATRILLE P., LIGHTNING J., LLOYD C., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SIMS M., SMALDON N., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,  
RA SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K.,  
RA WATERSON R., WATSON A., WEINSTOCK L., WILKINSON-SPROAT J.,  
RA WOHLDMAN P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL NATURE 368:32-38(1994).  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: L14429; G289772; -  
DR PIR: S44906; S44906.  
DR WORMPEP; ZK652.8; CE00453.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 194 AA; 23017 MW; F02A0951 CRC32;  
  
Query Match 0.8%; Score 7; DB 1; Length 194;  
Best Local Similarity 100.0%; Pred. No. 2.80e+00;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 93 ELEELSK 99  
QY 444 ELEELSK 450  
  
Search completed: Sat Aug 28 14:31:18 1999  
Job time : 67 secs.

**This Page Blank (uspto)**



\*\*\*\*\*

WIRELESS

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:31:37 1999; MasPar time 50.34 Seconds  
991.009 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-41

Description: (1-914) from US09049696.pep

Perfect Score: 914

Sequence: 1 MGPKSSVFILHLLEGAL.....GIHLKIMWKWIGELQLSLA 914

Scoring table: TABLE unitprotatable

Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database: sprenb19

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp\_unclassified  
13:sp-vertebrate 14:sp\_virus

Statistics: Mean 3.949; Variance 0.390; scale 10.133

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	40	4.4	913	11	GOB-5 PROTEIN.	4.75e-116
2	17	1.9	901	11	CHLORIDE CHANNEL CACC.	2.91e-29
3	13	1.4	109	6	UNKNOWN PROTEIN (FRAGM	3.64e-16
4	11	1.2	342	6	LU-ECAM-1.	3.67e-10
5	11	1.2	794	6	LU-ECAM-1.	3.67e-10
6	11	1.2	820	6	LU-ECAM-1.	3.67e-10
7	11	1.2	905	6	LU-ECAM-1.	3.67e-10
8	8	0.9	245	13	PROPIOMELANOTROPIN(PO	2.91e-02
9	8	0.9	391	2	MPT-SYNTHASE SULFURYL	2.91e-02
10	8	0.9	548	5	ALPHA ESTERASE.	2.91e-02
11	8	0.9	550	10	SAR DNA-BINDING PROTEI	2.91e-02
12	8	0.9	663	4	CYCLIN T2A.	2.91e-02
13	8	0.9	726	4	CYCLIN T1.	2.91e-02
14	8	0.9	730	4	CYCLIN T2B.	2.91e-02
15	8	0.9	730	4	CYCLIN T2B.	2.91e-02
16	7	0.8	63	2	HYPOTHETICAL 7.2 KD PR	4.55e+00
17	7	0.8	65	5	UBIQUITIN (FRAGMENT).	4.55e+00
18	7	0.8	76	5	UBIQUITIN (FRAGMENT).	4.55e+00
19	7	0.8	80	2	HYPOTHETICAL 9.2 KD PR	4.55e+00
20	7	0.8	83	14	KILLER VIRUS OF YEAST	4.55e+00

21	7	0.8	84	14	O90986	GP120 (FRAGMENT).	4.55e+00
22	7	0.8	86	14	P89730	ENVELOPE GLYCOPROTEIN	4.55e+00
23	7	0.8	107	2	P72164	HEAT SHOCK PROTEIN (FR	4.55e+00
24	7	0.8	112	11	Q60880	OLFACTORY RECEPTOR 4B	4.55e+00
25	7	0.8	123	1	P72356	HYPOTHETICAL 13.6 KD P	4.55e+00
26	7	0.8	127	11	P61927	GAG-MYB PROTEIN (FRAGM	4.55e+00
27	7	0.8	128	2	O67514	HYPOTHETICAL 15.4 KD P	4.55e+00
28	7	0.8	142	2	P95184	HYPOTHETICAL 15.9 KD P	4.55e+00
29	7	0.8	144	1	P95921	ORF C01023.	4.55e+00
30	7	0.8	157	5	O61165	CYSTINE PROTEASE (FRA	4.55e+00
31	7	0.8	160	14	O85700	B-TROPIC GAG REGION P3	4.55e+00
32	7	0.8	160	14	O85701	N-TROPIC GAG REGION P3	4.55e+00
33	7	0.8	170	2	P96325	PUTATIVE MINOR FIMBRIA	4.55e+00
34	7	0.8	172	5	O45439	F35C5.4 PROTEIN.	4.55e+00
35	7	0.8	175	1	O28131	ISOCHORISMATASE (ENTB)	4.55e+00
36	7	0.8	179	14	O83118	ORF 20.1 KDA.	4.55e+00
37	7	0.8	197	9	O38349	ORF198.	4.55e+00
38	7	0.8	212	5	O23041	COSMID T22B7.	4.55e+00
39	7	0.8	218	2	O51131	TRANSFERIN BINDING PR	4.55e+00
40	7	0.8	221	2	O51134	CYTIDYLATE KINASE (CMK	4.55e+00
41	7	0.8	244	2	O31614	YJBP PROTEIN.	4.55e+00
42	7	0.8	254	2	O33644	MINOR FIMBRIAL SUBUNIT	4.55e+00
43	7	0.8	260	2	P71102	RESTRICTION ENDONUCLEA	4.55e+00
44	7	0.8	261	1	O73936	HYPOTHETICAL 29.0 KD P	4.55e+00
45	7	0.8	263	14	O85554	GAG P30 PROTEIN.	4.55e+00
46	7	0.8	263	14	O80879	CA PROTEIN.	4.55e+00
47	7	0.8	269	2	O07015	HYPOTHETICAL 30.0 KD P	4.55e+00
48	7	0.8	271	10	O81793	HYPOTHETICAL 28.5 KD P	4.55e+00
49	7	0.8	287	1	O00409	METHYLVIOLOGEN-REDUCIN	4.55e+00
50	7	0.8	287	4	O99565	NK RECEPTOR.	4.55e+00
51	7	0.8	293	13	O91953	HYPOTHETICAL 32.5 KD P	4.55e+00
52	7	0.8	297	5	O27555	GLYCERALDEHYDE-3-PHOSP	4.55e+00
53	7	0.8	307	4	O14624	KILLER CELL RECEPTOR.	4.55e+00
54	7	0.8	325	4	O14623	KILLER CELL RECEPTOR.	4.55e+00
55	7	0.8	326	2	O05269	HYPOTHETICAL 35.8 KD P	4.55e+00
56	7	0.8	326	3	O05573	HYPOTHETICAL 36.6 KD P	4.55e+00
57	7	0.8	327	2	O25525	GMP REDUCTASE (GUAC).	4.55e+00
58	7	0.8	335	4	O43363	HXA1_HUMAN HOMEOBOX PR	4.55e+00
59	7	0.8	342	4	O14622	KILLER CELL RECEPTOR.	4.55e+00
60	7	0.8	342	4	O43803	NK RECEPTOR.	4.55e+00
61	7	0.8	351	2	O03390	PUTATIVE GONOCOCCAL SE	4.55e+00
62	7	0.8	351	2	O06205	HYPOTHETICAL 38.1 KD P	4.55e+00
63	7	0.8	359	5	O62433	Y41E3.3 PROTEIN.	4.55e+00
64	7	0.8	360	4	O99559	NK RECEPTOR.	4.55e+00
65	7	0.8	360	4	P78401	NK-RECEPTOR.	4.55e+00
66	7	0.8	365	5	O20846	CODED FOR BY C. ELEGAN	4.55e+00
67	7	0.8	373	4	O02833	PUTATIVE TRANSCRIPTION	4.55e+00
68	7	0.8	377	4	O99562	NK RECEPTOR.	4.55e+00
69	7	0.8	377	4	O99560	NK RECEPTOR.	4.55e+00
70	7	0.8	377	4	O14621	KILLER CELL RECEPTOR.	4.55e+00
71	7	0.8	377	4	O99706	NATURAL KILLER CELL RE	4.55e+00
72	7	0.8	377	4	O43534	NATURAL KILLER CELL IN	4.55e+00
73	7	0.8	382	4	O99564	NK RECEPTOR.	4.55e+00
74	7	0.8	382	4	O43469	NATURAL KILLER CELL RE	4.55e+00
75	7	0.8	382	4	P78403	NK-RECEPTOR.	4.55e+00
76	7	0.8	383	5	O17645	C34F6.4 PROTEIN.	4.55e+00
77	7	0.8	396	1	O26243	CONSERVED PROTEIN.	4.55e+00
78	7	0.8	400	3	O74916	PUTATIVE ACETYLORNITHI	4.55e+00
79	7	0.8	401	10	O64696	T31E10.15 PROTEIN.	4.55e+00
80	7	0.8	409	5	O22100	SIMILAR TO ACETOACETYL	4.55e+00
81	7	0.8	411	14	O99265	GLYCOPROTEIN G.	4.55e+00
82	7	0.8	425	2	P95472	COLR, COLS AND ORF222	4.55e+00
83	7	0.8	435	5	O45500	HYPOTHETICAL 49.4 KD P	4.55e+00
84	7	0.8	443	3	O60158	F39B2.8 PROTEIN.	4.55e+00
85	7	0.8	458	2	O34324	DIHYDROLIPONAMIDE DEHYD	4.55e+00
86	7	0.8	465	10	O23004	F6P2.11 PROTEIN.	4.55e+00
87	7	0.8	468	3	O59696	HMG-BOX CONTAINING TRA	4.55e+00
88	7	0.8	472	1	O27206	METHYL-VIOLOGEN-REDUCI	4.55e+00
89	7	0.8	472	1	O50783	METHYL-VIOLOGEN-REDUCI	4.55e+00
90	7	0.8	483	8	O36097	CYTOCHROME OXIDASE SUB	4.55e+00
91	7	0.8	526	3	O74133	ALK7.	4.55e+00
92	7	0.8	534	2	O31241	POLYSIN.	4.55e+00
93	7	0.8	536	11	O60588	ENDOGENOUS ECOTROPIC M	4.55e+00

94	7	0.8	536 14	Q85963	GAG...ENV {PROVIRUS}	4.55e+00	167	6	0.7	82 14	Q41045	A563L PROTEIN.	3.40e+02
95	7	0.8	536 14	Q83383	(CLONE RV-DBA/2-AGED)	4.55e+00	168	6	0.7	82 14	Q84761	FMDV RNA OF PRIMARY TR	3.40e+02
96	7	0.8	537 14	O56222	PR65.	4.55e+00	169	6	0.7	83 13	Q03659	ANTIFREEZE PROTEIN 6 (	3.40e+02
97	7	0.8	537 14	O67457	PR65.	4.55e+00	170	6	0.7	84 5	Q21640	COSMID R02E12.	3.40e+02
98	7	0.8	537 14	Q83397	GAG.	4.55e+00	171	6	0.7	86 14	Q70938	ENVELOPE GLYCOPROTEIN	3.40e+02
99	7	0.8	537 14	Q83361	GAG. POLYPROTEIN.	4.55e+00	172	6	0.7	88 14	Q39983	INFLUENZA A/EQUINE/PRA	3.40e+02
100	7	0.8	538 14	Q12373	POLYPROTEIN.	4.55e+00	173	6	0.7	90 5	Q44091	VACUOLAR ATPASE 14KD S	3.40e+02
101	7	0.8	538 14	Q41249	GAG POLYPROTEIN.	4.55e+00	174	6	0.7	91 2	Q24964	HYPOTHETICAL 10.7 KD P	3.40e+02
102	7	0.8	538 14	O85734	RETROVIRIDAE COMPLETE	4.55e+00	175	6	0.7	92 2	P74779	HYPOTHETICAL 9.7 KD PR	3.40e+02
103	7	0.8	539 2	O51679	CCON (EC 1.9.3.1).	4.55e+00	176	6	0.7	97 2	Q46404	CORRESPONDS TO A 97 KM	3.40e+02
104	7	0.8	539 2	O39736	GAG POLYPROTEIN.	4.55e+00	177	6	0.7	97 1	Q27930	HYPOTHETICAL 10.4 KD P	3.40e+02
105	7	0.8	549 10	Q42274	T CITOPLASM MALE STERI	4.55e+00	178	6	0.7	98 10	O65307	CAFFEOL-COENZYME A TR	3.40e+02
106	7	0.8	551 3	Q05180	PHOSPHOPROTEIN PHOSPHA	4.55e+00	179	6	0.7	99 2	O32775	HYPOTHETICAL 11.4 KD P	3.40e+02
107	7	0.8	553 14	O56223	COMPLETE GENOME.	4.55e+00	180	6	0.7	99 11	Q83308	LONG INTERSPERSED REPE	3.40e+02
108	7	0.8	554 11	O62592	FBR-MURINE OSTEOSARCOM	4.55e+00	181	6	0.7	100 14	Q72945	ENVELOPE GLYCOPROTEIN	3.40e+02
109	7	0.8	561 2	Q05316	HYPOTHETICAL 62.6 KD P	4.55e+00	182	6	0.7	102 2	O54771	HYPOTHETICAL 11.6 KD P	3.40e+02
110	7	0.8	565 3	O14335	HYPOTHETICAL 59.8 KD Z	4.55e+00	183	6	0.7	102 2	O86228	HYPOTHETICAL 11.5 KD P	3.40e+02
111	7	0.8	577 2	O33047	HYPOTHETICAL 61.5 KD P	4.55e+00	184	6	0.7	102 4	Q92840	ORF1 (FRAGMENT).	3.40e+02
112	7	0.8	586 13	P70077	MU-CALPAIN LARGE SUBUN	4.55e+00	185	6	0.7	104 14	O89540	ENVELOPE GLYCOPROTEIN	3.40e+02
113	7	0.8	597 14	Q76138	GP120 AND GP41 (FRAGME	4.55e+00	186	6	0.7	104 11	Q83144	OUTER MEMBRANE CYTOCHR	3.40e+02
114	7	0.8	602 5	O09432	PROBABLE NA(+)/H(+) AN	4.55e+00	187	6	0.7	104 2	Q48985	THIOREDOXIN (FRAGMENT)	3.40e+02
115	7	0.8	634 4	Q75114	KIAA0616 PROTEIN (FRAG	4.55e+00	188	6	0.7	105 2	O68851	HYPOTHETICAL 12.0 KD P	3.40e+02
116	7	0.8	644 14	Q77802	ENVELOPE POLYPROTEIN (	4.55e+00	189	6	0.7	107 5	O01413	SERINE PROTEASE INHIBI	3.40e+02
117	7	0.8	705 2	Q51286	(M978) TBP2 GENE (FRAG	4.55e+00	190	6	0.7	107 5	Q21377	COSMID K09C4.	3.40e+02
118	7	0.8	708 2	O05166	PCPA.	4.55e+00	191	6	0.7	107 8	P92540	ORF1076.	3.40e+02
119	7	0.8	713 11	O88666	CALPAIN I LARGE SUBUNI	4.55e+00	192	6	0.7	107 14	Q79510	(MALAWI PATIENT 12203)	3.40e+02
120	7	0.8	713 11	O35350	CALPAIN 1	4.55e+00	193	6	0.7	109 5	O17287	W10D9.5 PROTEIN.	3.40e+02
121	7	0.8	713 11	P97571	MU-CALPAIN LARGE SUBUN	4.55e+00	194	6	0.7	111 9	O64106	HYPOTHETICAL 12.5 KD P	3.40e+02
122	7	0.8	715 13	Q42133	MUCL (EC 3.4.22.17).	4.55e+00	195	6	0.7	111 2	O85708	THLC (FRAGMENT).	3.40e+02
123	7	0.8	734 4	O60916	CGMP-DEPENDANT PROTEIN	4.55e+00	196	6	0.7	112 2	O06774	HYPOTHETICAL 11.6 KD P	3.40e+02
124	7	0.8	736 3	O59672	PROBABLE ATP-DEPENDENT	4.55e+00	197	6	0.7	113 5	P91185	COSMID C53C11.	3.40e+02
125	7	0.8	746 6	Q29514	TRANSMEMBRANE PROTEIN	4.55e+00	198	6	0.7	114 2	O67784	HYPOTHETICAL 13.9 KD P	3.40e+02
126	7	0.8	762 4	O00125	CGMP-DEPENDENT PROTEIN	4.55e+00	199	6	0.7	119 14	Q65008	ORF3+13K.	3.40e+02
127	7	0.8	762 11	O43237	TYPE II CGMP-DEPENDENT	4.55e+00	200	6	0.7	119 5	O45637	K02B7.3 PROTEIN.	3.40e+02
128	7	0.8	762 11	O64595	CGMP-DEPENDENT PROTEIN	4.55e+00	201	6	0.7	120 13	Q91982	HYPOTHETICAL 13.5 KD P	3.40e+02
129	7	0.8	762 11	Q61410	PROTEIN KINASE, CGMP-D	4.55e+00	202	6	0.7	121 14	O56883	RNA DEPENDENT RNA POLY	3.40e+02
130	7	0.8	789 11	P70505	FERTILIN ALPHA.	4.55e+00	203	6	0.7	121 1	O73941	HYPOTHETICAL 13.3 KD P	3.40e+02
131	7	0.8	848 5	Q21489	M03C11.2 PROTEIN.	4.55e+00	204	6	0.7	123 3	P87282	YDR271CP.	3.40e+02
132	7	0.8	900 5	O44838	H19M22.2 PROTEIN.	4.55e+00	205	6	0.7	123 5	Q18672	HYPOTHETICAL PROTEIN C	3.40e+02
133	7	0.8	932 11	O35802	INTER-ALPHA-INHIBITOR	4.55e+00	206	6	0.7	125 14	O90829	ENVELOPE PROTEIN (FRAG	3.40e+02
134	7	0.8	981 6	P79100	CAPCITATIVE CALCIUM E	4.55e+00	207	6	0.7	128 14	O89454	ORF (CEL06).	3.40e+02
135	7	0.8	996 10	O24293	CHLOROPLAST INNER ENVE	4.55e+00	208	6	0.7	131 3	Q07790	ORF YDR053W.	3.40e+02
136	7	0.8	996 10	O24303	IAP100.	4.55e+00	209	6	0.7	131 1	O46256	CONSERVED PROTEIN.	3.40e+02
137	7	0.8	1123 5	O18431	D2085.5 PROTEIN.	4.55e+00	210	6	0.7	131 1	O29082	IRON-SULFUR CLUSTER BI	3.40e+02
138	7	0.8	1136 1	O58384	1136AA LONG HYPOTHETIC	4.55e+00	211	6	0.7	131 4	Q16553	IRG-E PRECURSOR.	3.40e+02
139	7	0.8	1187 5	O27533	W08D2.5 PROTEIN.	4.55e+00	212	6	0.7	132 2	O53178	VERY HYPOTHETICAL PROT	3.40e+02
140	7	0.8	1259 10	O23102	KINESIN-LIKE PROTEIN.	4.55e+00	213	6	0.7	134 9	Q38304	L9	3.40e+02
141	7	0.8	1259 10	O23226	KINESIN-LIKE CALMODULI	4.55e+00	214	6	0.7	135 14	O37070	POLYPROTEIN (FRAGMENT)	3.40e+02
142	7	0.8	1261 10	Q39130	CALMODULIN-BINDING PRO	4.55e+00	215	6	0.7	139 5	O17694	C53A5.1 PROTEIN.	3.40e+02
143	7	0.8	1263 4	O15044	KIAA0335.	4.55e+00	216	6	0.7	145 2	Q52215	REPLICATION PROTEIN.	3.40e+02
144	7	0.8	1496 4	O92626	MYELOBLAST KIAA0230 (F	4.55e+00	217	6	0.7	145 2	O05353	SIMILAR TO GROEL PROTE	3.40e+02
145	7	0.8	1734 11	P70355	GAG-POL POLYPROTEIN.	4.55e+00	218	6	0.7	146 5	O17504	MULTIPROTEIN BRIDGING	3.40e+02
146	7	0.8	1736 14	O92808	PR190.	4.55e+00	219	6	0.7	149 5	O17648	COSCL2.5 PROTEIN.	3.40e+02
147	7	0.8	1736 14	Q83362	GAG-POL POLYPROTEIN.	4.55e+00	220	6	0.7	150 14	O11389	STRAIN RHPV-A L1 (FRAG	3.40e+02
148	7	0.8	1738 14	O39735	GAG-POL POLYPROTEIN (P	4.55e+00	221	6	0.7	150 2	P75336	POTATIVE LIPOPROTEIN.	3.40e+02
149	7	0.8	2157 3	Q00333	RAD9.	4.55e+00	222	6	0.7	151 13	O93230	CDIC2 (FRAGMENT).	3.40e+02
150	7	0.8	2261 5	O62175	F15D3.9 PROTEIN.	4.55e+00	223	6	0.7	151 2	Q51329	DNA FOR NARB GENE AND	3.40e+02
151	7	0.8	2292 14	O66850	COMPLETE RNA GENOME PR	4.55e+00	224	6	0.7	152 14	O12679	HO3 PARTIAL PUTATIVE O	3.40e+02
152	7	0.8	2292 14	O66765	POLYPROTEIN, COMPLETE	4.55e+00	225	6	0.7	152 2	P75213	W03 ORF152 PROTEIN.	3.40e+02
153	7	0.8	2292 14	Q89272	COMPLETE VIRAL PROTEIN	4.55e+00	226	6	0.7	154 5	O76857	BCL7-LIKE PROTEIN.	3.40e+02
154	7	0.8	2292 14	O66763	PRP8 PROTEIN.	4.55e+00	227	6	0.7	155 5	O77038	SERINE PROTEINASE (FRA	3.40e+02
155	7	0.8	2403 5	O15881	PRP8 PROTEIN HOMOLOGUE	4.55e+00	228	6	0.7	158 5	O02430	CYTOSOLIC CU-ZN SUPERO	3.40e+02
156	7	0.8	3587 2	O30408	TYROCIDINE SYNTHETASE	4.55e+00	229	6	0.7	158 14	O71028	DUTPASE (EC 3.6.1.23)	3.40e+02
157	7	0.8	4910 3	O12019	SIMILARITY NEAR N-TERM	4.55e+00	230	6	0.7	158 2	O67534	HYPOTHETICAL 17.3 KD P	3.40e+02
158	6	0.7	5355 2	O68006	BACITRACIN SYNTHETASE	4.55e+00	231	6	0.7	162 1	O26526	HYPOTHETICAL 18.6 KD P	3.40e+02
159	6	0.7	22 10	Q39606	NIT-1 GENE ENCODING NI	3.40e+02	232	6	0.7	164 2	P96708	YDGI PROTEIN.	3.40e+02
160	6	0.7	47 2	O50804	HYPOTHETICAL 5.4 KD PR	3.40e+02	233	6	0.7	165 11	Q93878	CALCIUM BINDING POLYPE	3.40e+02
161	6	0.7	62 2	O05907	DINITROGENASE BETA SUB	3.40e+02	234	6	0.7	166 3	Q08004	CHROMOSOME XII READING	3.40e+02
162	6	0.7	68 8	Q34771	ATPASE SUBUNIT 8.	3.40e+02	235	6	0.7	166 4	Q16123	CARDIAC VENTRICULAR MY	3.40e+02
163	6	0.7	69 14	O84751	FMDV GENES FOR THE STR	3.40e+02	236	6	0.7	166 4	Q14908	CARDIAC VENTRICULAR MY	3.40e+02
164	6	0.7	76 2	O32409	LIGHT HARVESTING 1 BET	3.40e+02	237	6	0.7	167 2	Q49533	P120 (FRAGMENT).	3.40e+02
165	6	0.7	78 14	Q98440	GENOME, PARTIAL SEQUEN	3.40e+02	238	6	0.7	168 5	O01414	SERINE PROTEASE INHIBI	3.40e+02
166	6	0.7	80 1	O26182	HYPOTHETICAL 9.2 KD PR	3.40e+02	239	6	0.7	168 2	O53035	HYPOTHETICAL 18.6 KD P	3.40e+02

240	6	0.7	172	2	033883	METHANOL DEHYDROGENASE	3.40e+02	313	248	10	023235	HYPOTHETICAL 29.5 KD P	3.40e+02
241	6	0.7	173	14	065439	ORF L2.	3.40e+02	314	249	5	P91170	SIMILARITY TO TOXOCARA	3.40e+02
242	6	0.7	173	14	027691	C2 PROTEIN.	3.40e+02	315	250	1	058975	250AA LONG HYPOTHETICA	3.40e+02
243	6	0.7	174	1	028724	HYPOTHETICAL 19.8 KD P	3.40e+02	316	250	1	P71848	HYPOTHETICAL 27.4 KD P	3.40e+02
244	6	0.7	177	14	065897	(ISOLATE GERMAN BYV-G)	3.40e+02	317	251	2	067940	HYPOTHETICAL 28.3 KD P	3.40e+02
245	6	0.7	178	2	P72564	M PROTEIN (FRAGMENT).	3.40e+02	318	251	5	018436	TRYPSIN-LIKE PROTEASE	3.40e+02
246	6	0.7	178	14	091881	ENVELOPE GLYCOPROTEIN	3.40e+02	319	254	5	076954	TRYPSIN PRECURSOR.	3.40e+02
247	6	0.7	178	2	P73780	TRANSPORASE.	3.40e+02	320	254	5	018434	TRYPSIN-LIKE PROTEASE	3.40e+02
248	6	0.7	179	11	063811	PROTEIN PHOSPHATASE 3.	3.40e+02	321	254	14	Q98241	MC073R.	3.40e+02
249	6	0.7	180	10	046332	PLASMID GLE4293, COMPL	3.40e+02	322	254	11	P70636	LAMININ CHAIN (FRAGMEN	3.40e+02
250	6	0.7	180	10	099204	HYPOTHETICAL 19.7 KD P	3.40e+02	323	256	10	Q07306	MYOSIN HEAVY CHAIN (FR	3.40e+02
251	6	0.7	181	1	048309	FLAGELLIN GENE 3'REGIO	3.40e+02	324	260	2	051395	STAGE 0 SPOULATION PR	3.40e+02
252	6	0.7	182	1	068704	YOPK (YOPQ) HOMOLOG (Y	3.40e+02	325	261	1	026555	HYPOTHETICAL 29.1 KD P	3.40e+02
253	6	0.7	182	10	023060	SIMILARITY TO IMMUNOGL	3.40e+02	326	261	2	054509	LYSR HOMOLOG (FRAGMENT	3.40e+02
254	6	0.7	183	14	092291	POLYPROTEIN (FRAGMENT)	3.40e+02	327	263	2	054487	SCCB PRECURSOR.	3.40e+02
255	6	0.7	183	2	069158	PDXJ (FRAGMENT).	3.40e+02	328	263	2	Q45546	DIVIB.	3.40e+02
256	6	0.7	184	14	091887	ENVELOPE GLYCOPROTEIN	3.40e+02	329	263	2	023222	CYTOLYTIC TOXIN.	3.40e+02
257	6	0.7	184	10	040446	TUMOR-RELATED PROTEIN	3.40e+02	330	263	2	005565	PHOSPHOLIPASE C (EC 3.	3.40e+02
258	6	0.7	184	1	030048	CYTOCHROME C OXIDASE,	3.40e+02	331	264	5	023569	ZK669-2 PROTEIN.	3.40e+02
259	6	0.7	186	4	035416	CTG26 (FRAGMENT).	3.40e+02	332	264	10	Q49121	RESISTANCE GENE HOMOLO	3.40e+02
260	6	0.7	190	10	082604	T2L5.6 PROTEIN.	3.40e+02	333	264	5	Q94813	UBIQUITIN (FRAGMENT).	3.40e+02
261	6	0.7	191	1	026250	PHENYLACRYLIC ACID DEC	3.40e+02	334	264	5	P73856	HYPOTHETICAL 29.6 KD P	3.40e+02
262	6	0.7	192	2	Q59510	MG(2+) TRANSPORT ATPAS	3.40e+02	335	265	10	Q41787	LIGHT HARVESTING CHLO	3.40e+02
263	6	0.7	193	5	021970	R13.4 PROTEIN.	3.40e+02	336	267	2	025642	HYPOTHETICAL 30.9 KD P	3.40e+02
264	6	0.7	193	3	042623	SPOULATION PROTEIN.	3.40e+02	337	270	2	033697	BIFUNCTIONAL ALDOLASE-	3.40e+02
265	6	0.7	195	14	072088	ENVELOPE GLYCOPROTEIN	3.40e+02	338	270	2	034649	YTLD.	3.40e+02
266	6	0.7	195	10	048671	ATRERIB.	3.40e+02	339	272	2	P71746	HYPOTHETICAL 29.3 KD P	3.40e+02
267	6	0.7	196	5	044415	T07A9.1 PROTEIN.	3.40e+02	340	272	2	034094	HEMOLYSIS-ASSOCIATED P	3.40e+02
268	6	0.7	197	2	005575	HYPOTHETICAL 21.4 KD P	3.40e+02	341	273	5	022007	SIMILAR TO TRANSPOSABL	3.40e+02
269	6	0.7	197	14	072093	ENVELOPE GLYCOPROTEIN	3.40e+02	342	274	2	066224	MANNOYL TRANSFERASE.	3.40e+02
270	6	0.7	198	2	049008	DNA POLYMERASE III (AL	3.40e+02	343	274	14	071104	33 KDA PROTEIN.	3.40e+02
271	6	0.7	199	5	076575	K07D4.4 PROTEIN.	3.40e+02	344	274	2	070052	MANNOYL TRANSFERASE.	3.40e+02
272	6	0.7	200	14	072095	ENVELOPE GLYCOPROTEIN	3.40e+02	345	274	3	Q06152	SIMILARITY TO RETROVIR	3.40e+02
273	6	0.7	202	14	073979	ENVELOPE GLYCOPROTEIN,	3.40e+02	346	276	2	053232	ENOYL-COA HYDRATASE.	3.40e+02
274	6	0.7	203	14	091255	COAT PROTEIN.	3.40e+02	347	276	2	P72988	UDP-3'-O-ACYL N-ACETVIG	3.40e+02
275	6	0.7	203	2	006370	REPRESSOR PROTEIN OF S	3.40e+02	348	277	5	044811	H04J21.2 PROTEIN.	3.40e+02
276	6	0.7	203	2	005578	HYPOTHETICAL 22.2 KD P	3.40e+02	349	278	2	086901	HYDROLASE, DXNB PROTEI	3.40e+02
277	6	0.7	205	3	034224	PUTATIVE RHO GDP-DISSO	3.40e+02	350	279	2	P74979	HEMIN BINDING PROTEIN.	3.40e+02
278	6	0.7	206	2	P54833	H10_ORF206 PROTEIN.	3.40e+02	351	279	10	080843	F4118.26 PROTEIN.	3.40e+02
279	6	0.7	207	14	072074	ENVELOPE GLYCOPROTEIN	3.40e+02	352	284	11	Q63759	HEPATIC GLYCOCEN-BINDI	3.40e+02
280	6	0.7	207	3	P87309	HYPOTHETICAL 23.2 KD P	3.40e+02	353	284	2	033239	HYPOTHETICAL 32.3 KD P	3.40e+02
281	6	0.7	208	10	032308	HOMOTIC PROTEIN HOMOL	3.40e+02	354	289	2	P73593	HYPOTHETICAL 32.0 KD P	3.40e+02
282	6	0.7	209	5	001400	BLASTODERM-SPECIFIC PR	3.40e+02	355	291	2	085307	PIPB.	3.40e+02
283	6	0.7	210	10	P93378	TUMOR-RELATED PROTEIN.	3.40e+02	356	293	2	066582	HYPOTHETICAL 33.0 KD P	3.40e+02
284	6	0.7	212	2	Q53596	PYRROLIDONE CARBOXYL P	3.40e+02	357	294	3	050052	GERANYLGERANYL TRANSFE	3.40e+02
285	6	0.7	217	6	077719	CYSTEINE-RICH SECRETOR	3.40e+02	358	295	11	008862	SECRETED FRIZZLED REL	3.40e+02
286	6	0.7	217	5	Q18969	D2013.1 PROTEIN.	3.40e+02	359	295	11	035297	SECRETED APOPTOSIS REL	3.40e+02
287	6	0.7	220	5	020008	F35B12.4 PROTEIN.	3.40e+02	360	295	1	026691	CONSERVED PROTEIN.	3.40e+02
288	6	0.7	223	14	092292	POLYPROTEIN (FRAGMENT)	3.40e+02	361	297	11	035929	RAS-LIKE GTP-BINDING P	3.40e+02
289	6	0.7	223	14	Q67797	PVII1 PROTEIN.	3.40e+02	362	299	2	006040	EPSL.	3.40e+02
290	6	0.7	223	2	P73431	RNA POLYMERASE SIGMA-E	3.40e+02	363	303	4	Q43753	NK RECEPTOR PRECURSOR.	3.40e+02
291	6	0.7	225	2	Q44479	ORF225.	3.40e+02	364	303	2	087968	ORF5.	3.40e+02
292	6	0.7	227	10	080562	F6E13.5 PROTEIN.	3.40e+02	365	303	2	066882	TRANSCRIPTIONAL REGULA	3.40e+02
293	6	0.7	227	6	P70960	HYPOTHETICAL 24.5 KD P	3.40e+02	366	304	4	Q43754	NK RECEPTOR PRECURSOR.	3.40e+02
294	6	0.7	227	6	002692	NEURONAL AXONAL MEMBRA	3.40e+02	367	304	2	070018	HYPOTHETICAL 32.8 KD P	3.40e+02
295	6	0.7	228	2	Q5819	ABC TRANSPORTER, ATP-B	3.40e+02	368	304	3	Q52184	ASPERGILLUS AWAMORI DN	3.40e+02
296	6	0.7	228	3	034105	PUTATIVE RIBULOSE-PHOS	3.40e+02	369	305	14	P88951	ORF 60.	3.40e+02
297	6	0.7	229	2	Q51648	IS401 TRANSPORASE SUBU	3.40e+02	370	305	5	015844	L1439.3.	3.40e+02
298	6	0.7	230	14	089283	NONSTRUCTURAL PROTEIN	3.40e+02	371	307	10	Q04951	SERINE/THREONINE PROTE	3.40e+02
299	6	0.7	230	14	041655	NONSTRUCTURAL PROTEIN	3.40e+02	372	307	2	084976	PHAN.	3.40e+02
300	6	0.7	231	1	Q59535	231AA LONG HYPOTHETICA	3.40e+02	373	309	13	Q93605	VITELLOGENIN (FRAGMENT	3.40e+02
301	6	0.7	231	10	Q40382	RNASE NE PRECURSOR.	3.40e+02	374	311	5	002348	MO2B1.1 PROTEIN.	3.40e+02
302	6	0.7	234	4	Q43404	CD30L PROTEIN.	3.40e+02	375	313	10	Q39885	POLYGALACTURONASE-INHI	3.40e+02
303	6	0.7	236	2	066597	HYPOTHETICAL 25.9 KD P	3.40e+02	376	314	5	Q18285	HYPOTHETICAL 36.6 KD P	3.40e+02
304	6	0.7	240	14	P87593	P346, 5A-KDA RNA DEPEN	3.40e+02	377	315	5	Q22488	SIMILAR TO C. ELEGANS	3.40e+02
305	6	0.7	243	2	Q32099	YUED PROTEIN.	3.40e+02	378	315	5	Q16012	IMMUNOGLOBULIN G BINDI	3.40e+02
306	6	0.7	244	1	Q59212	CONSERVED HYPOTHETICAL	3.40e+02	379	318	10	Q05957	PHOSPHONATE BIOSYNTHES	3.40e+02
307	6	0.7	245	1	Q28187	HYPOTHETICAL 28.6 KD P	3.40e+02	380	319	2	Q31543	YFQJ PROTEIN.	3.40e+02
308	6	0.7	245	10	Q42984	ACC SYNTHASE (EC 4.4.1	3.40e+02	381	319	2	069544	PUTATIVE SUGAR DEHYDRAT	3.40e+02
309	6	0.7	246	14	Q69103	UL28 PROTEIN (FRAGMENT	3.40e+02	382	321	2	Q30541	ACCB.	3.40e+02
310	6	0.7	247	5	Q18308	ZK909.1 PROTEIN.	3.40e+02	383	322	2	P73168	HYPOTHETICAL 35.6 KD P	3.40e+02
311	6	0.7	247	13	Q42608	TRYPSINOGEN A3 PRECURS	3.40e+02	384	324	5	045337	F10A3.8 PROTEIN.	3.40e+02
312	6	0.7	248	1	027838	CONSERVED PROTEIN.	3.40e+02	385	325	1	036955	325AA LONG HYPOTHETICA	3.40e+02

386	6	0.7	326	2	Q55915	HYPOTHETICAL 36.3 KD P	3.40e+02	459	6	0.7	333	2	Q55580	NA+/H+ ANTIporter.	3.40e+02
387	6	0.7	328	1	Q26443	RHAMNOSYL TRANSFERASE.	3.40e+02	460	6	0.7	384	1	Q27075	N-ETHYLAMMELINE CHLORO	3.40e+02
388	6	0.7	328	2	Q66891	FLAGELLAR SWITCH PROTE	3.40e+02	461	6	0.7	384	2	Q60164	PROTON-TRANSLOCATING N	3.40e+02
389	6	0.7	330	9	Q64310	AL-PROTEIN.	3.40e+02	462	6	0.7	385	5	Q47474	ROGA10.2 PROTEIN.	3.40e+02
390	6	0.7	330	5	Q16403	C36C5.10 PROTEIN.	3.40e+02	463	6	0.7	385	1	Q58116	385AA LONG HYPOTHETICA	3.40e+02
391	6	0.7	330	2	Q84570	HYPOTHETICAL 38.0 KD P	3.40e+02	464	6	0.7	386	13	Q73860	OVALBUMIN.	3.40e+02
392	6	0.7	330	5	Q16399	C36C5.6 PROTEIN.	3.40e+02	465	6	0.7	388	14	Q98189	MC021L.	3.40e+02
393	6	0.7	332	14	Q12345	OUTER CAPSID PROTEIN.	3.40e+02	466	6	0.7	388	9	Q38198	UNKNOWN PROTEIN.	3.40e+02
394	6	0.7	332	14	Q67543	OUTER CAPSID GLYCOPROT	3.40e+02	467	6	0.7	330	2	Q34228	ORF43X9 PROTEIN.	3.40e+02
395	6	0.7	332	14	Q67540	OUTER CAPSID GLYCOPROT	3.40e+02	468	6	0.7	330	2	Q87153	ORF22-14 PROTEIN.	3.40e+02
396	6	0.7	332	14	Q67542	OUTER CAPSID GLYCOPROT	3.40e+02	469	6	0.7	332	1	Q30312	CONSERVED HYPOTHETICAL	3.40e+02
397	6	0.7	332	14	Q12344	OUTER CAPSID PROTEIN.	3.40e+02	470	6	0.7	332	14	Q98571	GLN-RICH.	3.40e+02
398	6	0.7	332	10	Q38909	XYLOGLUCAN ENDOTRANSGL	3.40e+02	471	6	0.7	333	5	Q02236	EO1G4.1 PROTEIN.	3.40e+02
399	6	0.7	332	14	Q84715	OUTER CAPSID PROTEIN.	3.40e+02	472	6	0.7	334	5	Q02362	TO9E8.4 PROTEIN.	3.40e+02
400	6	0.7	334	5	Q17563	F07B10.3 PROTEIN.	3.40e+02	473	6	0.7	334	4	Q15361	ZINC FINGER PROTEIN.	3.40e+02
401	6	0.7	334	2	Q46553	INDOLEPYRUVATE DECARBO	3.40e+02	474	6	0.7	336	10	Q81749	HYPOTHETICAL 44.2 KD P	3.40e+02
402	6	0.7	335	2	Q46258	HYPOTHETICAL 37.9 KD P	3.40e+02	475	6	0.7	337	2	Q85687	HOOK ASSOCIATED PROTEI	3.40e+02
403	6	0.7	335	1	Q57746	335AA LONG HYPOTHETICA	3.40e+02	476	6	0.7	339	2	Q53510	SERINE-THREONINE PROTE	3.40e+02
404	6	0.7	336	10	Q23776	ACTIN 5 (FRAGMENT).	3.40e+02	477	6	0.7	400	2	Q31200	IMIDAZOLONE PROPIONATE	3.40e+02
405	6	0.7	336	2	Q05633	HYPOTHETICAL 36.5 KD P	3.40e+02	478	6	0.7	401	3	Q14243	HYPOTHETICAL 46.4 KD P	3.40e+02
406	6	0.7	337	2	Q96539	ORE26.	3.40e+02	479	6	0.7	402	10	Q23560	HYPOTHETICAL 45.5 KD P	3.40e+02
407	6	0.7	337	4	Q14868	EFFECTOR CELL PROTEASE	3.40e+02	480	6	0.7	405	2	P72553	CPS3M.	3.40e+02
408	6	0.7	337	5	Q01436	SERINE PROTEASE INHIBI	3.40e+02	481	6	0.7	405	10	Q65263	FEN23.19 PROTEIN.	3.40e+02
409	6	0.7	338	2	Q33090	HYPOTHETICAL 36.7 KD P	3.40e+02	482	6	0.7	405	2	P72521	HYPOTHETICAL 44.1 KD P	3.40e+02
410	6	0.7	339	2	Q56136	SPAT, SIPB, SIPC, SIPD	3.40e+02	483	6	0.7	406	14	Q90757	ANK2.	3.40e+02
411	6	0.7	340	2	Q87991	PUTATIVE MEMBRANE PROT	3.40e+02	484	6	0.7	406	2	Q56466	L-TRFA (PLASMIID R751 T	3.40e+02
412	6	0.7	340	2	Q69647	PUTATIVE ANION TRANSPO	3.40e+02	485	6	0.7	407	2	Q44365	TRAB.	3.40e+02
413	6	0.7	341	4	Q14945	NKAT2B.	3.40e+02	486	6	0.7	407	2	P71176	TRAF1.	3.40e+02
414	6	0.7	341	4	Q43472	NATURAL KILLER CELL IN	3.40e+02	487	6	0.7	409	10	Q40582	SUP2 (FRAGMENT).	3.40e+02
415	6	0.7	341	4	Q14944	NKAT2A.	3.40e+02	488	6	0.7	409	2	Q55598	HYPOTHETICAL 44.9 KD P	3.40e+02
416	6	0.7	342	3	Q74289	GERANYLGERANYL DIPHOSP	3.40e+02	489	6	0.7	410	2	Q52386	CHLOROBENZENE DIOXIGEN	3.40e+02
417	6	0.7	343	5	Q23636	C31A11.8 PROTEIN.	3.40e+02	490	6	0.7	410	14	Q69973	GP120 (FRAGMENT).	3.40e+02
418	6	0.7	343	5	Q17298	ZK250.4 PROTEIN.	3.40e+02	491	6	0.7	411	10	Q02763	PUTATIVE BZIP-LIKE DNA	3.40e+02
419	6	0.7	344	5	Q02102	COSMID C18E3.	3.40e+02	492	6	0.7	411	13	Q90306	REGENERATION INDUCED 2	3.40e+02
420	6	0.7	344	10	Q81781	HYPOTHETICAL 39.1 KD P	3.40e+02	493	6	0.7	412	14	Q69969	GP120 (FRAGMENT).	3.40e+02
421	6	0.7	344	8	Q79914	NADH DEHYDROGENASE SUB	3.40e+02	494	6	0.7	412	3	Q43029	FRUCTOSYL AMINE.	3.40e+02
422	6	0.7	345	5	Q44916	W10G11.5 PROTEIN.	3.40e+02	495	6	0.7	414	14	P89954	CORE, ENV AND PART OF	3.40e+02
423	6	0.7	346	10	Q41641	FERRIDOXIN-NADP+ REDUC	3.40e+02	496	6	0.7	414	14	Q89241	UL43 PROTEIN.	3.40e+02
424	6	0.7	348	4	Q14951	NKAT6.	3.40e+02	497	6	0.7	414	5	Q46150	(SEA URCHIN).	3.40e+02
425	6	0.7	349	5	Q45849	T27C5.5 PROTEIN.	3.40e+02	498	6	0.7	414	8	Q37174	GLYCERALDEHYDE-3-PHOSP	3.40e+02
426	6	0.7	349	2	Q07146	HYPOTHETICAL 37.9 KD P	3.40e+02	499	6	0.7	415	14	Q77798	ENVELOPE POLYPROTEIN (	3.40e+02
427	6	0.7	349	4	Q39018	BULLOUS EMPHIGOID AUT	3.40e+02	500	6	0.7	415	5	Q47123	F49D11.4 PROTEIN.	3.40e+02
428	6	0.7	350	10	Q96513	PEROXIDASE ATP5A (EC 1	3.40e+02	501	6	0.7	416	3	Q74411	PUTATIVE CELL WALL BIO	3.40e+02
429	6	0.7	350	8	Q33572	KINETOPLAST APOCYTOCHR	3.40e+02	502	6	0.7	417	5	Q02080	SIMILAR TO CCAAT/ENHAN	3.40e+02
430	6	0.7	355	4	Q14318	FK-506 BINDING PROTEIN	3.40e+02	503	6	0.7	417	6	Q29038	TENASCIN-X (FRAGMENT).	3.40e+02
431	6	0.7	355	11	Q35465	FK-506 BINDING PROTEIN	3.40e+02	504	6	0.7	418	4	Q60235	AIRWAY TRYPSIN-LIKE PR	3.40e+02
432	6	0.7	357	10	Q42927	NARINGENIN 3-DIOXYGENA	3.40e+02	505	6	0.7	420	1	Q28025	CONSERVED HYPOTHETICAL	3.40e+02
433	6	0.7	357	10	Q49982	FLAVANONE 3-HYDROXYLAS	3.40e+02	506	6	0.7	422	2	Q52191	ENOLASE (EC 4.2.1.11)	3.40e+02
434	6	0.7	358	13	Q57588	PUTATIVE TRANSCRIPTION	3.40e+02	507	6	0.7	422	3	Q74547	HYPOTHETICAL 46.9 KD P	3.40e+02
435	6	0.7	358	10	Q80411	PHARBITIS KNITD-LIKE	3.40e+02	508	6	0.7	423	1	Q27300	HOMOSERINE DEHYDROGENA	3.40e+02
436	6	0.7	362	5	Q16146	MOM-2.	3.40e+02	509	6	0.7	423	2	Q51724	PHLE.	3.40e+02
437	6	0.7	363	5	Q93193	C07A4.3 PROTEIN.	3.40e+02	510	6	0.7	423	2	P95746	HYPOTHETICAL 44.1 KD P	3.40e+02
438	6	0.7	364	2	P73423	CBID PROTEIN.	3.40e+02	511	6	0.7	424	2	Q47628	TIEB PROTEIN.	3.40e+02
439	6	0.7	364	10	Q24329	HYPOTHETICAL 40.0 KD P	3.40e+02	512	6	0.7	427	14	Q68466	POLYPROTEIN (FRAGMENT)	3.40e+02
440	6	0.7	365	14	Q66736	GAG PROTEIN (FRAGMENT)	3.40e+02	513	6	0.7	427	14	Q68463	POLYPROTEIN (FRAGMENT)	3.40e+02
441	6	0.7	365	3	Q00104	ENDOPOLYGALACTURONASE	3.40e+02	514	6	0.7	427	14	Q68487	POLYPROTEIN (FRAGMENT)	3.40e+02
442	6	0.7	369	1	P95879	ORF C06032.	3.40e+02	515	6	0.7	427	14	Q68493	POLYPROTEIN (FRAGMENT)	3.40e+02
443	6	0.7	371	10	Q04399	A-TYPE CYCLIN.	3.40e+02	516	6	0.7	427	14	Q68485	POLYPROTEIN (FRAGMENT)	3.40e+02
444	6	0.7	372	10	Q04112	FLAVANONE 3-HYDROXYLAS	3.40e+02	517	6	0.7	427	14	Q68491	POLYPROTEIN (FRAGMENT)	3.40e+02
445	6	0.7	373	14	Q90697	GLYCOPROTEIN E (FRAGME	3.40e+02	518	6	0.7	427	14	Q68489	POLYPROTEIN (FRAGMENT)	3.40e+02
446	6	0.7	374	10	Q43741	NARINGENIN 3-DIOXYGENA	3.40e+02	519	6	0.7	427	14	Q68484	POLYPROTEIN (FRAGMENT)	3.40e+02
447	6	0.7	375	10	Q43557	UNIDENTIFIED PRECURSOR	3.40e+02	520	6	0.7	427	14	Q68486	POLYPROTEIN (FRAGMENT)	3.40e+02
448	6	0.7	375	1	Q30258	COENZYME PQQ SYNTHESIS	3.40e+02	521	6	0.7	427	14	Q68492	POLYPROTEIN (FRAGMENT)	3.40e+02
449	6	0.7	376	5	Q19380	COSMID F13B9.	3.40e+02	522	6	0.7	427	2	Q83835	NITROGEN FIXATION PROT	3.40e+02
450	6	0.7	377	10	Q82497	T12H20.1 PROTEIN.	3.40e+02	523	6	0.7	428	14	Q67812	EARLY REGION 3 GENES.	3.40e+02
451	6	0.7	378	5	Q16946	APCRB2.	3.40e+02	524	6	0.7	428	2	Q69759	2-PHOSPHONOPROPIONATE	3.40e+02
452	6	0.7	378	3	Q12076	SAGA.	3.40e+02	525	6	0.7	431	5	Q46230	MYOSIN HEAVY CHAIN (FR	3.40e+02
453	6	0.7	379	5	Q27192	UBIQUITIN.	3.40e+02	526	6	0.7	432	1	Q59545	432AA LONG HYPOTHETICA	3.40e+02
454	6	0.7	379	5	Q01878	SIMILARITY TO DROSOPHI	3.40e+02	527	6	0.7	434	2	Q51900	FERRIC ENTEROBACTIN ES	3.40e+02
455	6	0.7	379	5	Q27191	UBIQUITIN.	3.40e+02	528	6	0.7	434	11	Q35317	PBX3A.	3.40e+02
456	6	0.7	380	10	Q43568	DNA-BINDING PROTEIN (F	3.40e+02	529	6	0.7	434	1	Q29087	HYPOTHETICAL 49.1 KD P	3.40e+02
457	6	0.7	381	10	Q39522	S-ADENOSYL-L-METHIONIN	3.40e+02	530	6	0.7	436	10	Q04389	B-LIKE CYCLIN.	3.40e+02
458	6	0.7	382	2	Q85187	SODIUM-PROTON ANTIPORT	3.40e+02	531	6	0.7	437	4	Q00534	BREAST CANCER SUPPRESS	3.40e+02

532	6	0.7	438	7	Q95366	NK RECEPTOR.	3.40e+02	605	533	2	054186	PUTATIVE TRANSPORT PRO	3.40e+02
533	6	0.7	440	2	P72990	HYPOTHETICAL 49.2 KD P	3.40e+02	606	535	2	052716	RIBITOL KINASE	3.40e+02
534	6	0.7	441	1	Q27107	HYPOTHETICAL 48.2 KD P	3.40e+02	607	537	5	077201	SERENDIPITY ALPHA PROT	3.40e+02
535	6	0.7	442	14	Q55759	PUTATIVE HELICASE.	3.40e+02	608	543	9	Q77924	DI3 OF T5 (FRAGMENT).	3.40e+02
536	6	0.7	444	2	Q85080	4-CHLOROBENZOATE TRANS	3.40e+02	609	545	2	025223	ADENINE SPECIFIC DNA M	3.40e+02
537	6	0.7	447	2	Q48675	NISK.	3.40e+02	610	546	4	015049	KIAA0341 (FRAGMENT).	3.40e+02
538	6	0.7	448	5	Q22372	T10B10.5 PROTEIN.	3.40e+02	611	548	5	Q32379	COSMID 2C395.	3.40e+02
539	6	0.7	451	2	P71003	HYPOTHETICAL 49.4 KD P	3.40e+02	612	548	2	Q32923	TPP-REQUIRING ENZYME.	3.40e+02
540	6	0.7	453	2	Q34772	YDC.	3.40e+02	613	551	5	044607	R09B5.4 PROTEIN.	3.40e+02
541	6	0.7	455	2	Q84398	HYPOTHETICAL 50.6 KD P	3.40e+02	614	551	6	Q28645	PODCALYXIN-LIKE PROTE	3.40e+02
542	6	0.7	455	4	Q14948	NKAT4B.	3.40e+02	615	551	4	015389	OB BINDING PROTEIN-2.	3.40e+02
543	6	0.7	455	4	Q14947	NKAT4A.	3.40e+02	616	552	14	Q92482	CHITINASE-CHI-A-ACNMPV	3.40e+02
544	6	0.7	457	4	Q14824	METALLOPROTEINASE PREC	3.40e+02	617	558	5	Q25687	APICAL MEMBRANE ANTIGE	3.40e+02
545	6	0.7	458	1	Q26298	CONSERVED PROTEIN.	3.40e+02	618	562	5	Q76382	C24G6.4 PROTEIN.	3.40e+02
546	6	0.7	459	11	Q70374	TCF-4 PROTEIN.	3.40e+02	619	562	5	Q44412	T07A9.10 PROTEIN.	3.40e+02
547	6	0.7	461	5	Q61125	KINASE RESPONSIVE TO S	3.40e+02	620	563	5	Q23857	VEGETATIVE STAGE. SPECI	3.40e+02
548	6	0.7	465	14	Q92450	LEF-4-ACNMPV ORF90.	3.40e+02	621	564	2	Q69188	C3-BINDING PROTEIN.	3.40e+02
549	6	0.7	466	2	Q85072	3-ISOPROPYLMALATE ISOM	3.40e+02	622	565	2	P95077	CYSG.	3.40e+02
550	6	0.7	467	5	Q18085	T13F3.5 PROTEIN.	3.40e+02	623	566	5	Q21740	R05D11.8 PROTEIN.	3.40e+02
551	6	0.7	469	14	P88838	NEURAMINIDASE.	3.40e+02	624	567	11	Q60505	PROVIRUS.	3.40e+02
552	6	0.7	469	14	P88837	NEURAMINIDASE.	3.40e+02	625	567	14	Q83496	CORONAVIRUS PROVIRAL D	3.40e+02
553	6	0.7	469	2	Q85065	ISOPROPYLMALATE ISOMER	3.40e+02	626	567	14	Q83496	CORONAVIRUS PROVIRAL D	3.40e+02
554	6	0.7	470	11	Q61382	TNF RECEPTOR ASSOCIATE	3.40e+02	627	567	11	Q70347	PORC-PI-1 GENE SIMILAR	3.40e+02
555	6	0.7	471	2	Q69486	PROTEIN EXPORT MEMBRAN	3.40e+02	628	569	10	Q49048	VACUOLAR PROTEIN SORTI	3.40e+02
556	6	0.7	471	10	Q39030	SERINE/THREONINE-PROTE	3.40e+02	629	569	5	Q93200	W04G3.8 PROTEIN.	3.40e+02
557	6	0.7	472	1	Q58606	472AA LONG HYPOTHETICA	3.40e+02	630	569	5	Q62071	C17D12.7 PROTEIN.	3.40e+02
558	6	0.7	473	3	Q12116	HYPOTHETICAL 52.2 KD P	3.40e+02	631	569	1	Q27599	TUNGSTEN FORMYLMETHANO	3.40e+02
559	6	0.7	474	14	Q89542	UL38.	3.40e+02	632	574	1	Q58327	HYPOTHETICAL PROTEIN M	3.40e+02
560	6	0.7	475	1	Q26479	O-ANTIGEN TRANSPORTER	3.40e+02	633	575	14	Q93506	GLYCOPROTEIN E.	3.40e+02
561	6	0.7	478	5	Q16571	F02E11.2 PROTEIN.	3.40e+02	634	576	5	Q24180	DEAF-1.	3.40e+02
562	6	0.7	478	1	Q29419	TRK POTASSIUM UPTAKE S	3.40e+02	635	581	14	Q02710	GAG POLYPROTEIN.	3.40e+02
563	6	0.7	478	1	Q28457	HYPOTHETICAL 51.8 KD P	3.40e+02	636	583	10	Q96476	LAP17.1A.	3.40e+02
564	6	0.7	479	2	Q83621	CONSERVED HYPOTHETICAL	3.40e+02	637	583	10	Q49973	KINASE ASSOCIATED PROT	3.40e+02
565	6	0.7	479	5	Q17880	C09H10.3 PROTEIN.	3.40e+02	638	584	5	Q02426	FA11.7 PROTEIN.	3.40e+02
566	6	0.7	479	2	Q59714	CATALASE (EC 1.11.1.6)	3.40e+02	639	584	5	Q02426	SNAIL HOMOLOG.	3.40e+02
567	6	0.7	480	4	Q75472	TUMOROUS IMAGINAL DISC	3.40e+02	640	588	14	Q83152	SIMILAR TO THE US22 GE	3.40e+02
568	6	0.7	482	2	Q53727	PPE FAMILY PROTEIN.	3.40e+02	641	590	4	Q43187	INTERLEUKIN-1 RECEPT	3.40e+02
569	6	0.7	488	2	Q87253	CONSERVED HYPOTHETICAL	3.40e+02	642	590	2	Q54525	67 KDA MYOSIN-CROSSREA	3.40e+02
570	6	0.7	488	10	Q22362	HYPOTHETICAL PROTEIN.	3.40e+02	643	591	11	Q07802	EARLY B-CELL FACTOR.	3.40e+02
571	6	0.7	491	1	Q27804	CONSERVED PROTEIN.	3.40e+02	644	591	14	Q03930	RNA-DEPENDENT RNA-POLY	3.40e+02
572	6	0.7	494	5	Q76941	GAGA FACTOR CLASS B-SP	3.40e+02	645	594	2	Q06398	HYPOTHETICAL 50.8 KD P	3.40e+02
573	6	0.7	495	14	Q89939	GAG PROTEIN.	3.40e+02	646	594	3	Q06761	MR56 PROTEIN.	3.40e+02
574	6	0.7	496	4	Q76005	PUTATIVE CA2+-TRANSPOR	3.40e+02	647	598	4	Q01652	HYPOTHETICAL 68.7 KD P	3.40e+02
575	6	0.7	496	14	Q72990	ENVELOPE GLYCOPROTEIN	3.40e+02	648	598	4	Q75268	F20887.1. PARTIAL CDS	3.40e+02
576	6	0.7	496	5	Q19639	F21D5.7 PROTEIN (FRAGM	3.40e+02	649	598	11	Q35865	ORPHAN NUCLEAR HORMONE	3.40e+02
577	6	0.7	499	5	Q44657	B0213.12 PROTEIN.	3.40e+02	650	601	2	Q68975	EXOPOLYGALACTURONASE.	3.40e+02
578	6	0.7	499	10	Q23173	HYPOTHETICAL 55.1 KD P	3.40e+02	651	601	11	Q62431	BRIGHT.	3.40e+02
579	6	0.7	501	2	Q05846	HYPOTHETICAL 54.0 KD P	3.40e+02	652	603	11	Q08919	NUMBLIKE.	3.40e+02
580	6	0.7	503	3	Q03034	D3703.2P.	3.40e+02	653	603	14	P88931	ORF 43.	3.40e+02
581	6	0.7	505	11	Q88808	TUBBY PROTEIN.	3.40e+02	654	607	5	Q02088	COSMID F43E2.	3.40e+02
582	6	0.7	505	10	Q64667	F22013.18.	3.40e+02	655	607	10	Q41792	CDPK-RELATED PROTEIN K	3.40e+02
583	6	0.7	506	2	Q53834	PHASE-2 FLAGELLIN STRU	3.40e+02	656	607	11	Q88486	CYTOPLASMIC DYNEIN INT	3.40e+02
584	6	0.7	506	3	Q00863	RAN1-LIKE PROTEIN KINA	3.40e+02	657	610	2	Q32582	TAXB.	3.40e+02
585	6	0.7	507	2	Q84150	MONOOXYGENASE.	3.40e+02	658	611	5	Q27063	STIALIDASE (FRAGMENT).	3.40e+02
586	6	0.7	509	3	Q59759	HYPOTHETICAL 57.8 KD P	3.40e+02	659	614	2	Q83498	ADENYLATE CYCLASE.	3.40e+02
587	6	0.7	510	10	Q40938	SUCROSE TRANSPORTER.	3.40e+02	660	616	1	Q26384	ATP-DEPENDENT CLP PROT	3.40e+02
588	6	0.7	511	4	Q75542	ABC TRANSPORTER (FRAGM	3.40e+02	661	619	2	Q54972	PNEUMOCOCCAL SURFACE P	3.40e+02
589	6	0.7	512	8	Q47022	PLASTID NDHB.	3.40e+02	662	619	2	Q68220	130KDA PROTEIN.	3.40e+02
590	6	0.7	512	14	Q66721	(GAG) GENE (POL) (FRAG	3.40e+02	663	622	2	P73946	HYPOTHETICAL 68.8 KD P	3.40e+02
591	6	0.7	514	6	Q62063	3', 5'-CYCLIC NUCLEOTI	3.40e+02	664	625	11	Q61468	MEGAKARYOCYTE POTENTIA	3.40e+02
592	6	0.7	514	11	Q62913	CD441 (FRAGMENT).	3.40e+02	665	626	5	Q46039	1-EVIDENCE-PREDICTED B	3.40e+02
593	6	0.7	515	2	Q68168	HSDM.	3.40e+02	666	629	2	Q86146	FLGE PROTEIN (FRAGMENT	3.40e+02
594	6	0.7	517	5	Q21292	K07F5.4 PROTEIN.	3.40e+02	667	632	1	Q50838	F48A11.5 PROTEIN.	3.40e+02
595	6	0.7	519	2	Q06919	NITROGEN FIXATION REGU	3.40e+02	668	634	1	Q05838	ORF634 (POT. 29 K PROT	3.40e+02
596	6	0.7	520	5	Q25397	CONOPRESSIN RECEPTOR 2	3.40e+02	669	634	5	Q41116	VIMAR.	3.40e+02
597	6	0.7	521	10	Q82507	F2P3.14 PROTEIN.	3.40e+02	670	635	2	Q45820	CSPB GENE.	3.40e+02
598	6	0.7	521	2	Q51603	LYSYL-TRNA SYNTHETASE.	3.40e+02	671	636	2	Q88039	PUTATIVE ABC TRANSPORT	3.40e+02
599	6	0.7	522	10	Q04948	PUTATIVE CYTOCHROME P4	3.40e+02	672	636	2	Q52200	PRIMASE.	3.40e+02
600	6	0.7	523	2	Q33753	SIGA BINDING PROTEIN P	3.40e+02	673	638	2	Q86147	FLGE PROTEIN (FRAGMENT	3.40e+02
601	6	0.7	523	14	Q69968	GLP20 (FRAGMENT).	3.40e+02	674	640	5	Q44954	C34B2.7 PROTEIN.	3.40e+02
602	6	0.7	524	3	Q74832	PUTATIVE MITOCHONDRIAL	3.40e+02	675	641	13	Q90851	GAMMAFBP-B.	3.40e+02
603	6	0.7	528	2	Q84352	ABC TRANSPORTER PROTEI	3.40e+02	676	641	4	Q15959	PML.	3.40e+02
604	6	0.7	529	14	Q85213	G PROTEIN.	3.40e+02	677	643	3	Q04177	D9509.17P.	3.40e+02

678	6	0.7	645	2	052683	FE-HYDROGENASE ALPHA S	3.40e+02	751	6	0.7	787	2	051272	CELL DIVISION PROTEIN,	3.40e+02
679	6	0.7	647	14	039278	COUNTERPART OF HSV-1 G	3.40e+02	752	6	0.7	788	11	035227	A DISINTEGRIN AND META	3.40e+02
680	6	0.7	647	14	077801	ENVELOPE POLYPROTEIN (	3.40e+02	753	6	0.7	788	5	018510	INSECT INTESTINAL MUCI	3.40e+02
681	6	0.7	649	10	081620	FRM12.9 PROTEIN	3.40e+02	754	6	0.7	791	2	025543	IRON-REGULATED OUTER M	3.40e+02
682	6	0.7	650	5	017866	F39H11.4 PROTEIN	3.40e+02	755	6	0.7	795	5	019870	F28C6.4 PROTEIN	3.40e+02
683	6	0.7	655	5	024937	TRANSHYDROGENASE (FRAG	3.40e+02	756	6	0.7	796	5	021258	PSI [ALTERNATIVELY SPL	3.40e+02
684	6	0.7	659	14	007019	GENOME POLYPROTEIN (CO	3.40e+02	757	6	0.7	797	4	015156	PWL-RAR PROTEIN	3.40e+02
685	6	0.7	659	14	066385	POLYPROTEIN (FRAGMENTS	3.40e+02	758	6	0.7	805	14	092431	ACMPV ORF66	3.40e+02
686	6	0.7	659	14	066372	POLYPROTEIN (FRAGMENTS	3.40e+02	759	6	0.7	805	5	018511	C44F1.5 PROTEIN (FRAGM	3.40e+02
687	6	0.7	659	14	066384	POLYPROTEIN (FRAGMENTS	3.40e+02	760	6	0.7	807	5	018511	INSECT INTESTINAL MUCI	3.40e+02
688	6	0.7	659	14	007018	GENOME POLYPROTEIN (CO	3.40e+02	761	6	0.7	809	5	090534	RSC12 (FRAGMENT)	3.40e+02
689	6	0.7	659	14	066391	POLYPROTEIN (FRAGMENTS	3.40e+02	762	6	0.7	822	14	096910	GLYCOPROTEIN B	3.40e+02
690	6	0.7	659	14	066378	POLYPROTEIN (FRAGMENTS	3.40e+02	763	6	0.7	822	2	033831	ENVELOPE GLYCOPROTEIN	3.40e+02
691	6	0.7	659	14	066377	POLYPROTEIN (FRAGMENTS	3.40e+02	764	6	0.7	822	14	056280	ENVELOPE GLYCOPROTEIN	3.40e+02
692	6	0.7	659	14	066373	POLYPROTEIN (FRAGMENTS	3.40e+02	765	6	0.7	824	4	016349	H-NUC	3.40e+02
693	6	0.7	659	14	066376	POLYPROTEIN (FRAGMENTS	3.40e+02	766	6	0.7	832	5	016461	F41B6.3 PROTEIN	3.40e+02
694	6	0.7	659	14	066375	POLYPROTEIN (FRAGMENTS	3.40e+02	767	6	0.7	834	5	019783	SIMILAR TO REPEARICH R	3.40e+02
695	6	0.7	659	2	059289	HEPARINASE III PROTEIN	3.40e+02	768	6	0.7	836	5	062201	F32B4.8 PROTEIN	3.40e+02
696	6	0.7	659	14	066383	POLYPROTEIN (FRAGMENTS	3.40e+02	769	6	0.7	838	10	065525	HYPOTHETICAL 92.8 KD P	3.40e+02
697	6	0.7	659	14	066388	POLYPROTEIN (FRAGMENTS	3.40e+02	770	6	0.7	840	3	005848	J1575L ORF	3.40e+02
698	6	0.7	661	14	093130	DBL PROTEIN	3.40e+02	771	6	0.7	847	10	048512	RECEPTOR KINASE 1	3.40e+02
699	6	0.7	663	14	098188	MC0191	3.40e+02	772	6	0.7	847	10	004383	SERINE /THREONINE KINA	3.40e+02
700	6	0.7	663	2	030874	CHOLINE BINDING PROTEI	3.40e+02	773	6	0.7	848	14	070012	ENVELOPE GLYCOPROTEIN	3.40e+02
701	6	0.7	665	13	090850	GAMMAFBP-A	3.40e+02	774	6	0.7	848	5	017061	F14F9.3 PROTEIN	3.40e+02
702	6	0.7	667	5	019367	F12F6.8 PROTEIN	3.40e+02	775	6	0.7	851	5	015892	UNCONVENTIONAL MYOSIN	3.40e+02
703	6	0.7	670	2	085825	FLAGELLIN SUBUNIT	3.40e+02	776	6	0.7	854	14	070198	ENVELOPE GLYCOPROTEIN	3.40e+02
704	6	0.7	676	13	090825	GAMMAFBP-C (FRAGMENT)	3.40e+02	777	6	0.7	855	14	073347	ENVELOPE GLYCOPROTEIN	3.40e+02
705	6	0.7	680	14	036361	TRANSPORT PROTEIN	3.40e+02	778	6	0.7	855	14	073344	ENVELOPE GLYCOPROTEIN	3.40e+02
706	6	0.7	684	5	024035	ENA POLYPEPTIDE	3.40e+02	779	6	0.7	856	14	012164	ENVELOPE GLYCOPROTEIN	3.40e+02
707	6	0.7	684	4	014050	ALPHA-3 TYPE IX COLLAG	3.40e+02	780	6	0.7	858	14	069535	STRUCTURAL PROTEIN	3.40e+02
708	6	0.7	689	2	051288	TRANSFERRIN-BINDING PR	3.40e+02	781	6	0.7	858	10	023647	STARCH BRANCHING ENZYM	3.40e+02
709	6	0.7	694	5	027549	ACETYL-COA SYNTHETASE	3.40e+02	782	6	0.7	859	2	027670	RNA POLYMERASE II, SEC	3.40e+02
710	6	0.7	697	3	074504	HYPOTHETICAL 77.8 KD P	3.40e+02	783	6	0.7	862	2	086144	FLAGELLAR HOOK SUBUNIT	3.40e+02
711	6	0.7	698	5	001809	CODED FOR BY C. ELEGAN	3.40e+02	784	6	0.7	868	10	043817	LIPOXYGENASE (EC 1.13.	3.40e+02
712	6	0.7	701	2	051403	NA+/H+ ANTIPORTER (NAP	3.40e+02	785	6	0.7	871	2	051904	(H14320)	3.40e+02
713	6	0.7	702	2	069749	TRANSFERIN BINDING PRO	3.40e+02	786	6	0.7	886	14	067652	CAPSID PROTEIN	3.40e+02
714	6	0.7	703	14	098650	COAT PROTEIN POLYPROTE	3.40e+02	787	6	0.7	888	2	085840	PYRUVATE PHOSPHATE DIK	3.40e+02
715	6	0.7	705	2	088022	PUTATIVE INTEGRAL MEMB	3.40e+02	788	6	0.7	902	10	081209	RESPIRATORY BURST OXID	3.40e+02
716	6	0.7	707	2	050198	TGNB-DEPENDENT RECEPTO	3.40e+02	789	6	0.7	904	4	008192	PROTODADHERIN 43 PRECU	3.40e+02
717	6	0.7	707	2	052307	XYLANASE	3.40e+02	790	6	0.7	904	1	054324	ORF904	3.40e+02
718	6	0.7	714	1	027427	PHOSPHORIBOSYLFORMYLGL	3.40e+02	791	6	0.7	907	5	026312	NONFIBRILLAR COLLAGEN	3.40e+02
719	6	0.7	715	5	002047	SIMILARITY TO RAT SH3	3.40e+02	792	6	0.7	915	2	065044	VAG8	3.40e+02
720	6	0.7	716	2	046317	ATP-DEPENDENT TRANSLOC	3.40e+02	793	6	0.7	917	1	051970	SIMILAR TO ESCHERICHIA	3.40e+02
721	6	0.7	722	5	001921	F23H11.8 PROTEIN	3.40e+02	794	6	0.7	918	2	024835	HYPOTHETICAL 93.4 KD P	3.40e+02
722	6	0.7	728	10	004512	SEQUENCE OF BAC F21M12	3.40e+02	795	6	0.7	921	11	061433	ALPHA-1 TYPE IX COLLAG	3.40e+02
723	6	0.7	729	2	069743	HYPOTHETICAL 74.5 KD P	3.40e+02	796	6	0.7	921	11	061940	PROCOLLAGEN, TYPE IX,	3.40e+02
724	6	0.7	730	10	082332	PUTATIVE TRANSPOSASE	3.40e+02	797	6	0.7	928	5	018696	C47G2.5 PROTEIN	3.40e+02
725	6	0.7	732	2	086424	FERRICHRONE IRON RECEPT	3.40e+02	798	6	0.7	931	14	087544	104K PROTEIN	3.40e+02
726	6	0.7	734	3	014122	HYPOTHETICAL 85.3 KD P	3.40e+02	799	6	0.7	932	13	057587	TENASCIN-W PRECURSOR	3.40e+02
727	6	0.7	735	10	081820	WALL-ASSOCIATED KINASE	3.40e+02	800	6	0.7	934	4	060622	PROTODADHERIN 43	3.40e+02
728	6	0.7	735	10	081714	HYPOTHETICAL 81.2 KD P	3.40e+02	801	6	0.7	940	13	093484	ALPHA 2 TYPE I COLLAGE	3.40e+02
729	6	0.7	736	11	061164	CCCTC-BINDING FACTOR (	3.40e+02	802	6	0.7	943	14	085665	GLYCOPROTEIN B	3.40e+02
730	6	0.7	743	11	054916	RALBP1 ASSOCIATED EPS	3.40e+02	803	6	0.7	944	6	079403	GLUCOSIDASE II	3.40e+02
731	6	0.7	745	10	048539	REOHAOSP (FRAGMENT)	3.40e+02	804	6	0.7	948	14	090050	GLYCOPROTEIN B HOMOLOG	3.40e+02
732	6	0.7	746	5	027526	T19B10.3 PROTEIN	3.40e+02	805	6	0.7	951	2	068078	ICE NUCLEATION PROTEIN	3.40e+02
733	6	0.7	748	10	022907	RNA HELICASE ISOLOG	3.40e+02	806	6	0.7	959	1	029498	HYPOTHETICAL 108.2 KD	3.40e+02
734	6	0.7	750	14	074105	GP160	3.40e+02	807	6	0.7	962	5	061360	CROL ALPHA	3.40e+02
735	6	0.7	751	4	075504	GAMMA2-ADAPTIN	3.40e+02	808	6	0.7	963	4	075185	KIAA0703 PROTEIN	3.40e+02
736	6	0.7	752	3	013400	TELOMERE-ASSOCIATED RE	3.40e+02	809	6	0.7	969	14	088656	VIRAL POLYPROTEIN (FRA	3.40e+02
737	6	0.7	753	4	009471	M-PHASE PHOSPHOPROTEIN	3.40e+02	810	6	0.7	975	4	043892	BICAUDAL-D	3.40e+02
738	6	0.7	754	2	053052	TRANSPOSASE	3.40e+02	811	6	0.7	979	10	093820	HYPOTHETICAL 110.0 KD	3.40e+02
739	6	0.7	757	3	013399	TELOMERE-ASSOCIATED RE	3.40e+02	812	6	0.7	1000	2	085086	LACTOFERRIN BINDING PR	3.40e+02
740	6	0.7	760	10	081732	HYPOTHETICAL 85.3 KD P	3.40e+02	813	6	0.7	1006	14	057191	DNA POLYMERASE (EC 2.7	3.40e+02
741	6	0.7	760	2	075376	PUTATIVE LIPOPROTEIN	3.40e+02	814	6	0.7	1008	2	049619	PUTATIVE 106.8 KD TRAN	3.40e+02
742	6	0.7	767	10	093410	VACUOLAR H+-PYROPHOSPH	3.40e+02	815	6	0.7	1009	11	062640	GLUTAMATE RECEPTOR DEL	3.40e+02
743	6	0.7	771	5	027338	KSR-1	3.40e+02	816	6	0.7	1012	4	043393	ATROPHIN-1 RELATED PRO	3.40e+02
744	6	0.7	776	5	009501	HYPOTHETICAL 89.0 KD P	3.40e+02	817	6	0.7	1012	4	045359	ATROPHIN-1 LIKE PROTEI	3.40e+02
745	6	0.7	781	14	082857	ENV PRECURSOR	3.40e+02	818	6	0.7	1013	6	004451	F21J9.13	3.40e+02
746	6	0.7	783	14	036393	HELICASE	3.40e+02	819	6	0.7	1014	6	062805	RETINAL ROD NA/CA-K EX	3.40e+02
747	6	0.7	785	3	075006	SUBUNIT OF THE FINAL S	3.40e+02	820	6	0.7	1033	5	024327	GP160-DTRK PRECURSOR	3.40e+02
748	6	0.7	785	4	075843	GAMMA2-ADAPTIN	3.40e+02	821	6	0.7	1039	2	P74552	HELICASE OF THE SNF2/R	3.40e+02
749	6	0.7	785	14	P89451	HERPES SIMPLEX VIRUS T	3.40e+02	822	6	0.7	1041	5	Q18031	NAD(P) TRANSHYDROGENAS	3.40e+02
750	6	0.7	786	5	016361	F44C8.2 PROTEIN	3.40e+02	823	6	0.7	1046	2	P96156	CHITODEXTRINASE	3.40e+02

824	6	0.7	1049	5	O00831	UNCONVENTIONAL MYOSIN	3.40e+02
825	6	0.7	1051	3	O60039	APSE GENE.	3.40e+02
826	6	0.7	1052	2	O65937	HYPOTHETICAL 104.6 KD	3.40e+02
827	6	0.7	1053	2	O08394	PROBABLE CYTOCHROME P4	3.40e+02
828	6	0.7	1071	14	P89047	POLYPROTEIN (FRAGMENT)	3.40e+02
829	6	0.7	1075	2	O06342	HYPOTHETICAL 115.0 KD	3.40e+02
830	6	0.7	1079	2	O49555	MEMBRANE PROTEIN.	3.40e+02
831	6	0.7	1083	5	O45604	TOIG1.3. PROTEIN.	3.40e+02
832	6	0.7	1085	2	O25577	CARBAMOYL-PHOSPHATE SY	3.40e+02
833	6	0.7	1086	5	O10039	PUTATIVE GUANYLATE CYC	3.40e+02
834	6	0.7	1087	5	O20948	F57C7.1B PROTEIN.	3.40e+02
835	6	0.7	1089	5	O19937	F31B12.2 PROTEIN.	3.40e+02
836	6	0.7	1089	14	O40947	ORF 73.	3.40e+02
837	6	0.7	1099	2	O54377	EXONUCLEASE REXB.	3.40e+02
838	6	0.7	1099	2	O49560	MHEP1.	3.40e+02
839	6	0.7	1101	5	O76369	F56H1.3 PROTEIN.	3.40e+02
840	6	0.7	1114	13	O42408	DELTAEP1.	3.40e+02
841	6	0.7	1116	2	P73507	DNA POLYMERASE III SUB	3.40e+02
842	6	0.7	1118	4	O92900	REGULATOR OF NONSENSE	3.40e+02
843	6	0.7	1127	13	O13046	AND-1 PROTEIN.	3.40e+02
844	6	0.7	1129	4	O00239	KIAA0221.	3.40e+02
845	6	0.7	1132	4	O14746	TELOMERASE REVERSE TRA	3.40e+02
846	6	0.7	1146	13	O90584	NOVEL COLLAGEN PROTEIN	3.40e+02
847	6	0.7	1162	14	O98148	ORF73 HOMOLOG.	3.40e+02
848	6	0.7	1164	14	O90027	HOMOLOG OF VACCINIA VI	3.40e+02
849	6	0.7	1170	4	Q16817	PLASMA MEMBRANE CALCIU	3.40e+02
850	6	0.7	1173	11	O63624	CTD-BINDING SR-LIKE PR	3.40e+02
851	6	0.7	1173	5	O94314	CODED FOR BY C. ELEGAN	3.40e+02
852	6	0.7	1201	5	O16844	KINESIN-RELATED PROTEI	3.40e+02
853	6	0.7	1203	5	O77312	MAL3P1.9 PROTEIN.	3.40e+02
854	6	0.7	1203	5	O22484	SIMILARITY TO RNA MET	3.40e+02
855	6	0.7	1209	11	O88904	HOMEODOMAIN-INTERACTI	3.40e+02
856	6	0.7	1212	4	O60305	KIAA0559 PROTEIN (FRAG	3.40e+02
857	6	0.7	1237	2	O84549	DNA POL III ALPHA.	3.40e+02
858	6	0.7	1240	13	P79773	INSULIN RECEPTOR SUBST	3.40e+02
859	6	0.7	1250	5	O20947	F57C7.1A PROTEIN.	3.40e+02
860	6	0.7	1254	5	O62559	SDMR2.	3.40e+02
861	6	0.7	1256	5	Q19103	HYPOTHETICAL PROTEIN F	3.40e+02
862	6	0.7	1265	10	O41460	KINESIN HEAVY CHAIN-LI	3.40e+02
863	6	0.7	1268	4	O75046	KIAA0458 PROTEIN.	3.40e+02
864	6	0.7	1279	2	O50314	PROTOPORPHYRIN IX MG C	3.40e+02
865	6	0.7	1285	11	O70305	SPINOCEREBELLAR ATAXIA	3.40e+02
866	6	0.7	1289	2	O34986	YVNB.	3.40e+02
867	6	0.7	1291	11	Q61812	PROTEIN TYROSINE PHOSP	3.40e+02
868	6	0.7	1295	2	P74881	PHOSPHORIBOSYLFORMYLGL	3.40e+02
869	6	0.7	1299	4	O60343	KIAA0603 PROTEIN.	3.40e+02
870	6	0.7	1302	5	O61542	NEUROGLIAN.	3.40e+02
871	6	0.7	1305	2	O06521	AGGREGATION SUBSTANCE.	3.40e+02
872	6	0.7	1311	4	O14376	HSGAK.	3.40e+02
873	6	0.7	1323	11	Q63381	NMDA RECEPTOR SUBUNIT	3.40e+02
874	6	0.7	1331	10	O49549	HYPOTHETICAL 148.1 KD	3.40e+02
875	6	0.7	1345	14	O66675	TEGUMENT PROTEIN.	3.40e+02
876	6	0.7	1368	2	O86094	RNA POLYMERASE B-SUBUN	3.40e+02
877	6	0.7	1371	2	P73337	SENSORY TRANSDUCTION H	3.40e+02
878	6	0.7	1386	4	O75064	KIAA0476 PROTEIN.	3.40e+02
879	6	0.7	1391	11	O55103	L-PERITAXIN.	3.40e+02
880	6	0.7	1400	2	O84454	PROTEIN EXPORT.	3.40e+02
881	6	0.7	1401	4	O15021	ORF.	3.40e+02
882	6	0.7	1458	3	Q13457	CUT4+.	3.40e+02
883	6	0.7	1462	2	O69383	GLUCOSYLTRANSFERASE-S.	3.40e+02
884	6	0.7	1462	2	O69392	GLUCOSYLTRANSFERASE-S.	3.40e+02
885	6	0.7	1462	2	O69389	GLUCOSYLTRANSFERASE-S	3.40e+02
886	6	0.7	1463	11	O70320	PHOSPHOLIPASE B.	3.40e+02
887	6	0.7	1483	5	O22528	T16G12.5 PROTEIN.	3.40e+02
888	6	0.7	1489	2	O53559	PGRS-FAMILY PROTEIN.	3.40e+02
889	6	0.7	1503	10	O22774	PUTATIVE CHLOROPLAST O	3.40e+02
890	6	0.7	1508	2	O52274	GLUCOSYLTRANSFERASE (E	3.40e+02
891	6	0.7	1518	5	O21442	LAMININ A (FRAGMENT).	3.40e+02
892	6	0.7	1518	2	O00600	GLUCOSYLTRANSFERASE I	3.40e+02
893	6	0.7	1520	4	O15087	KIAA0383 (FRAGMENT).	3.40e+02
894	6	0.7	1532	13	O90894	190 KD TENASCIN PRECUR	3.40e+02
895	6	0.7	1532	4	O02802	BULLOUS EMPHYGMOID AUT	3.40e+02
896	6	0.7	1539	5	Q19611	SIMILARITY TO PUTATIVE	3.40e+02

897	6	0.7	1544	11	O54989	SULFONYLUREA RECEPTOR	3.40e+02
898	6	0.7	1567	11	O35243	ANTIGEN CONTAINING EPI	3.40e+02
899	6	0.7	1576	5	P90841	F20G4.1 PROTEIN.	3.40e+02
900	6	0.7	1577	2	O55265	GLUCOSYLTRANSFERASE PR	3.40e+02
901	6	0.7	1582	11	P70532	SULFONYLUREA RECEPTOR.	3.40e+02
902	6	0.7	1594	2	O60146	PEPTIDE SYNTHETASE (FR	3.40e+02
903	6	0.7	1612	2	O53020	168 KD SURFACE-LAYER P	3.40e+02
904	6	0.7	1620	4	O99996	YOTIAO (FRAGMENT).	3.40e+02
905	6	0.7	1642	4	O14869	YOTIAO.	3.40e+02
906	6	0.7	1645	5	O27448	GLUTAMINE-DEPENDENT CA	3.40e+02
907	6	0.7	1680	10	O64588	T29F13.11 PROTEIN.	3.40e+02
908	6	0.7	1684	10	O80815	TRF5.22 PROTEIN.	3.40e+02
909	6	0.7	1704	2	O51816	ARG-GINGIPAIN-1. PROTEI	3.40e+02
910	6	0.7	1706	2	O51838	PROTEASE PRECURSOR.	3.40e+02
911	6	0.7	1706	3	O42900	PROTEIN KINASE.	3.40e+02
912	6	0.7	1707	14	O96806	PUTATIVE REPLICASE.	3.40e+02
913	6	0.7	1707	14	O67704	PUTATIVE REPLICASE.	3.40e+02
914	6	0.7	1714	13	O90995	200 KD TENASCIN PRECUR	3.40e+02
915	6	0.7	1802	5	O17163	A2 (IV) BASEMENT MEMBR	3.40e+02
916	6	0.7	1805	14	O41809	POLYPROTEIN.	3.40e+02
917	6	0.7	1810	13	O90824	CYTOTACTIN 200KD PRECU	3.40e+02
918	6	0.7	1856	4	O99407	ANKYRIN.	3.40e+02
919	6	0.7	1858	3	P78615	FATTY ACID SYNTHASE, A	3.40e+02
920	6	0.7	1873	14	O83044	METHYLTRANSFERASE.	3.40e+02
921	6	0.7	1883	5	O24376	CHROMODOMAIN-HELICASE-	3.40e+02
922	6	0.7	1893	5	O24279	CNO.	3.40e+02
923	6	0.7	1904	9	O38319	ORF1904.	3.40e+02
924	6	0.7	1943	11	O61307	ANKYRIN 3 (ANKYRIN G)	3.40e+02
925	6	0.7	1978	4	O15154	TRIP230.	3.40e+02
926	6	0.7	2090	5	O94247	CODED FOR BY C. ELEGAN	3.40e+02
927	6	0.7	2092	4	O14667	KTAA0100 PROTEIN.	3.40e+02
928	6	0.7	2109	14	O65867	237K PROTEIN.	3.40e+02
929	6	0.7	2115	14	O86500	M33 RNA FOR A NONSTRUC	3.40e+02
930	6	0.7	2116	14	O40955	NONSTRUCTURAL PROTEIN.	3.40e+02
931	6	0.7	2153	5	O94246	CODED FOR BY C. ELEGAN	3.40e+02
932	6	0.7	2161	14	O91724	RNA-DEPENDENT RNA POLY	3.40e+02
933	6	0.7	2161	14	O91901	REPLICASE.	3.40e+02
934	6	0.7	2193	14	O66849	COMPLETE GENOME.	3.40e+02
935	6	0.7	2203	14	O66577	POLYPROTEIN.	3.40e+02
936	6	0.7	2206	14	O83017	POLYPROTEIN 1A.	3.40e+02
937	6	0.7	2212	4	O93072	MYELOBLAST KIAA0246 PR	3.40e+02
938	6	0.7	2218	14	O09705	L PROTEIN.	3.40e+02
939	6	0.7	2226	14	O66502	POLYPROTEIN.	3.40e+02
940	6	0.7	2227	14	O66775	POLYPROTEIN.	3.40e+02
941	6	0.7	2247	14	O39818	ATP-BINDING CASSETTE T	3.40e+02
942	6	0.7	2273	4	O51112	ABCR.	3.40e+02
943	6	0.7	2273	4	O60915	ATP-BINDING CASSETTE T	3.40e+02
944	6	0.7	2273	4	P78363	ATP-BINDING CASSETTE T	3.40e+02
945	6	0.7	2310	11	O35800	PUTATIVE PHOSPHATIDYL	3.40e+02
946	6	0.7	2335	3	O14356	POLYPROTEIN.	3.40e+02
947	6	0.7	2344	14	O86119	POLYPROTEIN.	3.40e+02
948	6	0.7	2344	14	O86114	POLYPROTEIN.	3.40e+02
949	6	0.7	2344	14	O86117	(SD).	3.40e+02
950	6	0.7	2352	5	O61240	HRNOTCH PROTEIN.	3.40e+02
951	6	0.7	2374	5	O45377	F22G12.5 PROTEIN.	3.40e+02
952	6	0.7	2408	4	O92566	MYELOBLAST KIAA0279 (F	3.40e+02
953	6	0.7	2493	3	O60055	HYPOTHETICAL 285.2 KD	3.40e+02
954	6	0.7	2528	5	O19317	F10F2.1 PROTEIN.	3.40e+02
955	6	0.7	2529	5	O24605	ZINC FINGER PROTEIN.	3.40e+02
956	6	0.7	2533	5	O27183	ALPHA-51D-IMMOBILIZATI	3.40e+02
957	6	0.7	2541	5	O19663	SIMILAR TO NON-MUSCLE M	3.40e+02
958	6	0.7	2543	5	P90649	156D SURFACE ANTIGEN.	3.40e+02
959	6	0.7	2559	5	O44113	PUTATIVE GUANINE NUCLE	3.40e+02
960	6	0.7	2559	5	O44381	SHAR PEI/DRHOGEF2.	3.40e+02
961	6	0.7	2635	14	P88955	ORF 64.	3.40e+02
962	6	0.7	2635	14	O40942	ORF 64.	3.40e+02
963	6	0.7	2647	5	P90580	FCR3-VART11-1 PROTEIN	3.40e+02
964	6	0.7	3010	14	P90192	POLYPROTEIN.	3.40e+02
965	6	0.7	3010	14	P88803	POLYPROTEIN.	3.40e+02
966	6	0.7	3010	14	P90191	POLYPROTEIN.	3.40e+02
967	6	0.7	3010	14	P90195	POLYPROTEIN.	3.40e+02
968	6	0.7	3010	14	P90193	POLYPROTEIN.	3.40e+02
969	6	0.7	3010	14	P90194	POLYPROTEIN.	3.40e+02

```
970 6 0.7 3010 14 Q81541 POLYPROTEIN. 3.40e+02
971 6 0.7 3010 14 Q93016 POLYPROTEIN. 3.40e+02
972 6 0.7 3010 14 Q68533 POLYPROTEIN. 3.40e+02
973 6 0.7 3011 14 Q36608 POLYPROTEIN. 3.40e+02
974 6 0.7 3011 14 Q81754 POLYPROTEIN. 3.40e+02
975 6 0.7 3011 14 Q36609 POLYPROTEIN. 3.40e+02
976 6 0.7 3011 14 Q36510 POLYPROTEIN. 3.40e+02
977 6 0.7 3011 14 Q34653 GENOME POLYPROTEIN. 3.40e+02
978 6 0.7 3011 14 Q36579 POLYPROTEIN. 3.40e+02
979 6 0.7 3014 14 Q86614 POLYPROTEIN. 3.40e+02
980 6 0.7 3015 14 Q92532 POLYPROTEIN. 3.40e+02
981 6 0.7 3016 14 Q92531 POLYPROTEIN. 3.40e+02
982 6 0.7 3019 14 Q92529 POLYPROTEIN. 3.40e+02
983 6 0.7 3066 11 Q62388 ATAXIA TELANGIECTASIA 3.40e+02
984 6 0.7 3122 14 P89459 VERY LARGE TEGUMENT PR 3.40e+02
985 6 0.7 3169 2 Q82916 TOXIN B. 3.40e+02
986 6 0.7 3247 14 Q65553 UL36. 3.40e+02
987 6 0.7 3313 11 Q88278 MEGF2. 3.40e+02
988 6 0.7 3473 14 Q83034 POLYPROTEIN. 3.40e+02
989 6 0.7 3507 5 Q23587 CODED FOR BY C. ELEGAN 3.40e+02
990 6 0.7 3722 2 P94873 ALPHA-AMINOADIPYL-CYST 3.40e+02
991 6 0.7 3896 14 P87514 PESTIVIRUS POLYPROTEIN 3.40e+02
992 6 0.7 3942 11 Q88737 BASSOON. 3.40e+02
993 6 0.7 4135 6 Q18977 TENASCIN-X. 3.40e+02
994 6 0.7 4292 4 Q15141 POLYCYSTIC KIDNEY DISE 3.40e+02
995 6 0.7 4302 4 Q15140 POLYCYSTIC KIDNEY DISE 3.40e+02
996 6 0.7 4351 11 Q88277 MEGF1. 3.40e+02
997 6 0.7 5069 2 Q52789 RIFAMYCIN POLYKETIDE S 3.40e+02
998 6 0.7 5486 2 Q30409 TYROCIDINE SYNTHETASE 3.40e+02
999 6 0.7 15281 3 Q09164 CYCLOSPORIN SYNTHETASE 3.40e+02
1000 6 0.7 26926 4 Q10466 TITIN, HEART ISOFORM N 3.40e+02
```

## ALIGNMENTS

```
RESULT 1 PRELIMINARY: PRT: 913 AA.
ID Q88826
AC Q88826;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GDB-5 PROTEIN.
GN GDB-5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RA KOMIYA T., TANIGAWA Y., HIROHASHI S.;
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet
   cells in mice.";
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AB017156; D1034712; -.
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 4.4%; Score 40; DB 11; Length 913;
Best Local Similarity 100.0%; Pred. No. 4.75e-116;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 615 IROGASPILRASVTALIESVNGTKVTLELDNGAGADATK 654
QY |||||
614 IROGASPILRASVTALIESVNGTKVTLELDNGAGADATK 653

RESULT 2 PRELIMINARY: PRT: 901 AA.
ID Q88860
AC Q88860;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CHLORIDE CHANNEL CACC.
OS MUS MUSCULUS (MOUSE).
```

```
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RA ROMIO L., MUSANTE L., CINTI R., MORAN O., SERI M., GALIETTA L.J.V.;
RT "Characterization of a murine gene homologous to the bovine CaCC
   chloride channel.";
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF052746; G3560547; -.
SQ SEQUENCE 901 AA; 100039 MW; E52BF02E CRC32;

Query Match 1.9%; Score 17; DB 11; Length 901;
Best Local Similarity 100.0%; Pred. No. 2.91e-29;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 FVHEWAHLRWGVFDEYN 169
QY |||||
154 FVHEWAHLRWGVFDEYN 170
```

```
RESULT 3 PRELIMINARY: PRT: 109 AA.
ID Q29282
AC Q29282;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-MAY-1997 (TREMBLREL. 03, LAST ANNOTATION UPDATE)
DE UNKNOWN PROTEIN (FRAGMENT).
OS SUS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA WINTEROE A.K., FREDHOLM M., DAVIES W.;
RL MAMM. GENOME 7:509-517(1996).
DR EMBL; F15082; G972064; -.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 12048 MW; AF422625 CRC32;

Query Match 1.4%; Score 13; DB 6; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.64e-16;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 RIVCLVLDKSGSM 13
QY |||||
305 RIVCLVLDKSGSM 317
```

```
RESULT 4 PRELIMINARY: PRT: 342 AA.
ID Q18744
AC Q18744;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE LU-ECAM-1.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,
RA GOODWIN A., PAULI B.U.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF001264; G2623769; -.
SQ SEQUENCE 342 AA; 39037 MW; 0DF25F74 CRC32;
```

```
Query Match 1.2%; Score 11; DB 6; Length 342;
Best Local Similarity 100.0%; Pred. No. 3.67e-10;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 FVHEWAHLRWG 163
```



```
QY 154 FVHEWAHLRWG 164
|||||
RESULT 5
ID O18742 PRELIMINARY: PRT: 794 AA.
AC O18742:
DT 01-JAN-1998 (TREMREL. 05, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE LU-ECAM-1.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,
RA GOODWIN A., PAULI B.U.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF001262; G2623765; -.
SQ SEQUENCE 794 AA; 88509 MW; 2D450EEB CRC32;

Query Match 1.2%; Score 11; DB 6; Length 794;
Best Local Similarity 100.0%; Pred. No. 3.67e-10;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 FVHEWAHLRWG 163
|||||
QY 154 FVHEWAHLRWG 164

RESULT 6
ID O18743 PRELIMINARY: PRT: 820 AA.
AC O18743:
DT 01-JAN-1998 (TREMREL. 05, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE LU-ECAM-1.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,
RA GOODWIN A., PAULI B.U.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF001263; G2623767; -.
SQ SEQUENCE 820 AA; 91464 MW; 171ED21C CRC32;

Query Match 1.2%; Score 11; DB 6; Length 820;
Best Local Similarity 100.0%; Pred. No. 3.67e-10;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 FVHEWAHLRWG 163
|||||
QY 154 FVHEWAHLRWG 164

RESULT 7
ID O18741 PRELIMINARY: PRT: 905 AA.
AC O18741:
DT 01-JAN-1998 (TREMREL. 05, CREATED)
DT 01-JAN-1998 (TREMREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)
DE LU-ECAM-1.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVINAE; BOS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
```

```
RA ELBLE R.C., WIDOM J., GRUBER A.D., ABDEL-GHANY M., LEVINE R.,
RA GOODWIN A., PAULI B.U.;
RL SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF001261; G2623763; -.
DR PFAM: PF00041; fn3; 1.
SQ SEQUENCE 905 AA; 101005 MW; 86E65459 CRC32;

Query Match 1.2%; Score 11; DB 6; Length 905;
Best Local Similarity 100.0%; Pred. No. 3.67e-10;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 FVHEWAHLRWG 163
|||||
QY 154 FVHEWAHLRWG 164

RESULT 8
ID Q91259 PRELIMINARY: PRT: 245 AA.
AC Q91259:
DT 01-NOV-1996 (TREMREL. 01, CREATED)
DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMREL. 09, LAST ANNOTATION UPDATE)
DE PROPIOMELANOTROPIN(POM) PRECURSOR.
OS PETROMYZON MARINUS (SEA LAMPREY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; CEPHALASPIDOMORPHI;
OC PETROMYZONTIFORMES; PETROMYZONTIDAE; PETROMYZON.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95374508.
RA TAKAHASHI A., AMEMIYA Y., SARASHI M., SOWER S., KAWAUCHI H.;
RT "Melanotropin and corticotropin are encoded on two distinct genes in
RL the most early evolved extant vertebrate lamprey.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 213:490-498(1995).
DR EMBL: D55629; D1010134; -.
KW SIGNAL.
FT SIGNAL 1 39 POTENTIAL.
FT CHAIN 102 121 MSH-B.
FT CHAIN 188 206 MSH-A.
FT CHAIN 211 245 BETA-END.
FT SEQUENCE 245 AA; 26615 MW; 85C9BBAA CRC32;

Query Match 0.9%; Score 8; DB 13; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.91e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 75 DGSEIVLL 82
|||||
QY 403 DGSEIVLL 410

RESULT 9
ID O54307 PRELIMINARY: PRT: 391 AA.
AC O54307:
DT 01-JUN-1998 (TREMREL. 06, CREATED)
DT 01-JUN-1998 (TREMREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMREL. 07, LAST ANNOTATION UPDATE)
DE MPT-SYNTHASE SULFURYLASE.
GN MOEB.
OS SYNECHOCOCUS SP. (STRAIN PCC 7942).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCOCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC9742;
RA RUBIO L.M., FLORES E., HERRERO A.;
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: Y16560; E1258370; -.
SQ SEQUENCE 391 AA; 42421 MW; D6A955CD CRC32;

Query Match 0.9%; Score 8; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 2.91e-02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 366 AGITGNV 373
```

2

QY 637 TVTLELDD 644

RESULT 15  
ID O60583 PRELIMINARY: PRT: 730 AA.  
AC O60583;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE CYCLIN T2B.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98167917.  
RA PENG J.M., ZHU Y., MILTON J.T., PRICE D.H.;  
RT "Identification of multiple cyclin subunits of human P-TEFb.";  
RL GENES DEV. 12:755-762(1998).  
DR EMBL; AF048732; G2981200; ..  
SQ SEQUENCE 730 AA; 81045 MW; A2B74978 CRC32;

Query Match 0.9%; Score 8; DB 4; Length 730;  
Best Local Similarity 100.0%; Pred. No. 2.91e-02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 227 TVTLELDD 234

QY 637 TVTLELDD 644

Search completed: Sat Aug 28 14:33:12 1999  
Job time : 95 secs.

**This Page Blank (uspto)**





```

Best Local Similarity 23.5%; Pred. No. 5.99e+01;
Matches 19; Conservative 22; Mismatches 39; Indels 1; Gaps 1;

Db 173 ilpdgasqstnfpmpssstpvqqlppaatqtlqqlgemsgpmqqltqplqgvtsl 232
QY 409 LLTGDENTISGCFNEVKQSGAIHTVALGPSAAQLEELSKMTGGTQTYASDQVQNNGL 468

Db 233 fsqvggtgggnpa-deeaaqm 252
QY 469 IDAFGALSSGNGAVSQRSIQL 489

RESULT 5
ID W64338 standard; Protein; 368 AA.
AC W64338.
DE 09-NOV-1998 (first entry)
KW Mycobacterium tuberculosis antigen Tb37-FL.
KW Tuberculosis; infection; diagnosis; antigen; Tb37-FL.
OS Mycobacterium tuberculosis strain H37Rv.
FH Key Location/Qualifiers
FT Misc_difference 10
FT /note= "unidentified"
PN W09816645-A2.
PD 23-APR-1998.
PF 07-OCT-1997; U18214.
PR 13-MAR-1997; US-818111.
PR 11-OCT-1996; US-729622.
PA (CORI-) CORIXA CORP.
PI Campos-Neto A, Dillon DC, Houghton R, Lodes M,
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
PI WPI: 98-251292/22.
PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and diagnosis of tuberculosis
PS Example 3; Page 143-144; 250pp; English.
CC This polypeptide comprises Mycobacterium tuberculosis antigen
CC Tb37-FL. It is encoded by genomic DNA isolated from a M.
CC tuberculosis strain H37Rv genomic library using a probe from
CC clone Tb38-1 (see V44384). The invention relates to compositions
CC and methods for diagnosing tuberculosis. It provides polypeptides
CC (see W64291-W64379) comprising an antigenic portion of a soluble
CC M. tuberculosis antigen, or an immunogenic portion of an M.
CC tuberculosis antigen, as well as DNA sequences encoding such
CC polypeptides, recombinant expression vectors and transformed or
CC transfected host cells. Also claimed are methods and diagnostic
CC kits for detecting M. tuberculosis infection in a patient using
CC these polypeptides, antibodies or oligonucleotide probes and
CC primers, for the diagnosis of tuberculosis.
SQ Sequence 368 AA;

Query Match 1.6%; Score 100; DB 35; Length 368;
Best Local Similarity 23.5%; Pred. No. 5.99e+01;
Matches 19; Conservative 22; Mismatches 39; Indels 1; Gaps 1;

Db 173 ilpdgasqstnfpmpssstpvqqlppaatqtlqqlgemsgpmqqltqplqgvtsl 232
QY 409 LLTGDENTISGCFNEVKQSGAIHTVALGPSAAQLEELSKMTGGTQTYASDQVQNNGL 468

Db 233 fsqvggtgggnpa-deeaaqm 252
QY 469 IDAFGALSSGNGAVSQRSIQL 489

RESULT 6
ID W32384 standard; Protein; 368 AA.
AC W32384.
DE 13-JAN-1998 (first entry)
KW Mycobacterium tuberculosis antigen Tb37-FL.
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT Misc_difference 10

```

```

FT W09709429-A2.
PN 13-MAR-1997.
PF 30-AUG-1996; U14675.
PR 12-JUL-1996; US-680573.
PR 01-SEP-1995; US-523435.
PR 22-SEP-1995; US-532136.
PR 22-MAR-1996; US-620280.
PR 05-JUN-1996; US-658800.
PA (CORI-) CORIXA CORP.
PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
PI Twardzik DR, Vedvick TH;
PI WPI: 97-192904/17.
PT New immunogenic polypeptide(s) from soluble M. tuberculosis antigens
PT - useful for diagnosis of M. tuberculosis infection
PS Example 3; Page 159-161; 190pp; English
CC A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
CC its variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis
CC antigen, Tb37-FL. The immunogenic polypeptide can be used to diagnose
CC M.tuberculosis infection by forming complexes with specific
CC antibodies in the sample. Fragments of DNA encoding the immunogenic
CC polypeptide can be used as diagnostic primers or probes and agents
CC that bind to the antigen, especially monoclonal antibodies or
CC equivalent polyclonal antibodies, are also used for diagnosis.
SQ Sequence 368 AA;

Query Match 1.6%; Score 100; DB 25; Length 368;
Best Local Similarity 23.5%; Pred. No. 5.99e+01;
Matches 19; Conservative 22; Mismatches 39; Indels 1; Gaps 1;

Db 173 ilpdgasqstnfpmpssstpvqqlppaatqtlqqlgemsgpmqqltqplqgvtsl 232
QY 409 LLTGDENTISGCFNEVKQSGAIHTVALGPSAAQLEELSKMTGGTQTYASDQVQNNGL 468

Db 233 fsqvggtgggnpa-deeaaqm 252
QY 469 IDAFGALSSGNGAVSQRSIQL 489

RESULT 7
ID W32452 standard; Protein; 368 AA.
AC W32452.
DE 09-JAN-1998 (first entry)
KW Mycobacterium tuberculosis antigen Tb37-FL.
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
OS Mycobacterium tuberculosis.
FH Key Location/Qualifiers
FT Misc_difference 10
FT /note= "Any amino acid"
PN W09709428-A2.
PN 13-MAR-1997.
PF 30-AUG-1996; U14674.
PR 12-JUL-1996; US-680574.
PR 01-SEP-1995; US-523436.
PR 22-SEP-1995; US-533634.
PR 22-MAR-1996; US-620874.
PR 05-JUN-1996; US-659683.
PA (CORI-) CORIXA CORP.
PI Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
PI Twardzik DR, Vedvick TH;
PI WPI: 97-192903/17.
PT New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
PT useful in vaccines for prevention or treatment of tuberculosis, also
PT for diagnosis
PS Example 3; Page 146-147; 168pp; English.
CC A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
CC its variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis
CC antigen, Tb37-FL. The immunogenic protein, and fusion proteins

```

KW endoglucanase; Bacillus agaradhaerens; cellulose binding domain;  
Cel5A.  
KW

This represents a *Bacillus agaradhaerens* endoglucanase enzyme. This is used in the construction of enzyme hybrids for liquefaction of starch. The enzyme hybrids contain amino acid sequences of alpha-amylase linked to a cellulose binding domain (CBD). The starch is liquefied by treating in aqueous medium, with such an enzyme hybrid. A recombinant expression vector comprising a construct containing isolated DNA encoding enzyme hybrids with amylolytic activity, promoter and stop signals can be used to transform host cells for the production of the recombinant enzyme hybrids. The enzyme hybrids are useful in industrial starch processing especially for the production of sweeteners. Hybrid enzymes have altered affinity for substrate and increased activity, resulting in at least 10% reduced calcium ion dependence, reduced formation of Maillard reaction products and reduced effect of alpha-amylase on subsequent saccharification.



N-PSDB; 102139.  
polypeptide(s) comprising immunogenic part of Leishmania antigen -

```

Db      273 f 273
Qy      865 F 865

RESULT 15
ID      W70213 standard; Protein: 403 AA.
AC      W70213;
DE      Leishmania antigen (first entry)
DE      Leishmania antigen LbeIF4A protein.
KW      Leishmania antigen: immune response; infection detection; therapy;
KW      humoral response induction; cellular response induction; cancer;
KW      interleukin-12 production.
OS      Leishmania sp.
PN      WO9835045-A2.
PD      13-AUG-1998.
PF      12-FEB-1998; U03002.
PR      27-AUG-1997; US-920609.
PR      12-FEB-1997; US-798841.
PA      (CORI-) CORIXA CORP.
PI      Campos-Neto A, Dillon DC, Reed SG, Skeiky YA, Webb JR;
PI      WPT: 98-447242/38.
DR      N-PSDB; V47559.
PT      New immunogenic fragments of Leishmania antigens and related nucleic
PT      acid, vectors and host cells - are useful for diagnosis, prevention
PT      and treatment of leishmaniasis, also to induce production of
PT      interleukin-12 generally
PS      Claim 27; Page 85-87; 194pp; English.
CC      This sequence represents a Leishmania antigen (LAG) of the invention,
CC      designated LbeIF4A. Compositions and vaccines containing the protein
CC      used to generate a protective or therapeutic immune response against
CC      Leishmania species donavani, chagasi, infantum, major, amazonensis,
CC      braziliensis, panamensis, tropica or guayanensis. They can also be us
CC      to detect infection (in a skin test). The compositions induce a humor
CC      and/or cellular response, specifically of Th1 type, particularly
CC      including induction of interleukin-12 (IL-12) production. They may th
CC      be used more generally to treat any condition (e.g. bacterial, viral
CC      protozoal infection, or cancer) which responds to IL-12.
SQ      Sequence 403 AA;

Query Match      1.6%; Score 104; DB 35; Length 403;
Best Local Similarity 24.8%; Pred. No. 3.30e-01;
Matches 30; Conservative 35; Mismatches 46; Indels 10; Gaps

Db      155 pgrsvdvikgartsesrlvvldeademsqfaqiyeifrlpkdiqv-alfsatmp 2133
Qy      753 PQGQIT-LK-AEIHGGSLNLTWTPGDDVDHG-TAHKY-IIIRISITSLDRDKFNESLQ 808
      |||:::| : : : : || : : : : || : : : : || : : : : || : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      214 eevlelt-kkfmrdprrilvkresltlegikqffiaeveehkiddlmdiyetsiaqsvi 872
Qy      809 VNTALIPKCANSEEV-FLEKPNITFENGTDLFIAIQAVDKVD-LKS--EISNIARVSL 864
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      273 f 273
Qy      865 F 865

Search completed: Mon Aug 23 13:02:03 1999
Job time : 75 secs.

```

\*\*\*\*\*  
MWISREH  
\*\*\*\*\* (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 12:58:56 1999; MasPar time 39.91 Seconds  
917.693 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-049-696-41  
Description: (1-914) from US09049696.pep  
Perfect Score: 6397  
Sequence: 1 MGPFKSSVFILHLLEGAL.....GIHLKIMWKWIGELQLSIA 914

Scoring table: PAM 150  
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 53.072; Variance 103.195; scale 0.514

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	5070	79.3	913	3	gob-5 protein - Mouse	0.00e+00
2	3206	50.1	905	2	Lu-ECAM-1 protein - b	0.00e+00
3	2915	45.6	820	2	Lu-ECAM-1 protein - b	0.00e+00
4	2911	45.5	794	2	hypothetical protein	0.00e+00
5	174	2.7	420	2	hypothetical protein	4.67e-10
6	113	1.8	300	2	1-phosphofructokinase	2.93e-01
7	112	1.8	464	2	probable cystathionin	3.91e-01
8	114	1.8	540	2	heat shock protein 56	2.20e-01
9	118	1.8	575	2	hypothetical protein	6.78e-02
10	116	1.8	1848	2	cellulose-binding pro	1.23e-01
11	108	1.7	263	2	hypothetical protein	1.21e+00
12	106	1.7	332	2	probable integrase -	2.10e+00
13	108	1.7	451	2	hypothetical protein	1.21e+00
14	108	1.7	451	2	hypothetical protein	1.21e+00
15	111	1.7	561	2	protoporphyrin IX mag	5.20e-01
16	110	1.7	629	2	protoporphyrin IX mag	6.90e-01
17	106	1.7	1010	2	brkA prtein - Bordet	2.10e+00
18	109	1.7	2020	2	ABC-type transport pr	9.15e-01
19	100	1.6	128	2	hypothetical protein	1.05e+01
20	101	1.6	148	2	legumin beta chain -	8.05e+00
21	103	1.6	235	2	flagellar assembly pr	4.73e+00
22	101	1.6	235	2	flagellar protein fli	8.05e+00
23	102	1.6	247	2	hypothetical protein	6.18e+00

24	105	1.6	290	2	D69518	conserved hypohetica	2.76e+00
25	101	1.6	361	2	F64355	conserved hypohetica	8.05e+00
26	100	1.6	368	2	G70802	probable pPE protein	1.05e+01
27	104	1.6	403	2	T02861	EIF-4A protein - Leis	3.62e+00
28	105	1.6	411	2	C64052	NADH dehydrogenase (u	2.76e+00
29	100	1.6	489	2	E69369	2-isopropylmalate syn	1.05e+01
30	100	1.6	526	2	S46766	hypothetical protein	1.05e+01
31	104	1.6	663	2	B70460	excinuclease ABC (EC	3.62e+00
32	100	1.6	682	2	S40037	pilJ protein - Pseudo	1.05e+01
33	103	1.6	759	2	D70422	cellulose synthase ca	4.73e+00
34	103	1.6	788	1	JDVLHH	DNA-directed DNA poly	4.73e+00
35	102	1.6	815	2	A64048	aspartate kinase (EC	6.18e+00
36	100	1.6	852	2	T00994	hypothetical protein	1.05e+01
37	102	1.6	885	2	S30350	inter-alpha-trypsin i	6.18e+00
38	100	1.6	1018	2	I37246	contactin - human	1.05e+01
39	100	1.6	1018	2	A54744	contactin 1 precursor	1.05e+01
40	100	1.6	1018	2	JC4211	neural adhesion prote	1.05e+01
41	100	1.6	1020	2	S05944	neuronal cell surface	1.05e+01
42	100	1.6	1021	2	A57112	contactin precursor -	1.05e+01
43	104	1.6	1143	2	S73173	DNA-directed RNA poly	3.62e+00
44	102	1.6	1239	2	G02750	DNA-directed DNA poly	6.18e+00
45	102	1.6	3566	2	A40701	tenascin-X precursor	6.18e+00

ALIGNMENTS

RESULT 1  
ENTRY JG0168 #type complete  
TITLE gob-5 protein - Mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 18-Mar-1999 #sequence\_revision 18-Mar-1999 #text\_change 18-Mar-1999

ACCESSIONS JG0168  
REFERENCE JG0168  
#authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.  
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351  
#title Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice.

#accession JG0168  
#status preliminary  
##residues 1-913 #label KOM  
##cross-references DDBJ:AB016592

SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 79.3%; Score 5070; DB 3; Length 913;  
Best Local Similarity 75.3%; Pred. No. 0.00e+00;  
Matches 692; Conservative 133; Mismatches 84; Indels 7; Gaps 5;

Db	1	MESLSPVFLIIHLLEGVLSLSLIQNNNGYEGIVIAIDHDVPEDEALQHKDMVTQA	60
Qy	1	MGPFKSSVFILHLLEGALSLSLIQNNNGYEGIVVADPNVPDETLIQQLKDMVTQA	60
Db	61	SPYLFEATGKRFYKFNAILIPESWAKPEYTRPKLETAKNADLVSTSTPLGNDEPYTE	120
Qy	61	SLYLFEATGKRFYKFNAILIPETWTKADYVRPKLETAKNADLVVAESTPPGNDEPYTE	120
Db	121	HIGACEGKIRIHLTDFDLGKLLTOYQODRFTVHEWAFHRCVGVNEYNDEKFLSKG	180
Qy	121	OMGNCGEKGERIHLTDFDLGKLLTOYQODRFTVHEWAFHRCVGVNEYNDEKFLSKG	180
Db	181	KPAVPCSAITKGNVRKCGSGCITNGKCVIDRTGLYKDCNCFVPPDPHONKASIMF	240
Qy	181	RQAVPCSAITGTVNKKQCGGSCYTK-RCFNVKTVGLYEKCFEVLQSRQTEKASIMF	239
Db	241	QNINSWVEFCFEKHNQENAPDNQNCNLRSTWIEOESDFKQTPMTAQPPAPFTSL	300
Qy	240	AQHVDSIVEFCFEQHNKEAPNKQKNCNLRSTWIEOESDFKQTPMTQPPNPFTSL	299
Db	301	LQIGQIRIVCLVDKSGSMLNDRRLNRNMQASRLFLQTVQGSWGVMTFDSAAYVOSEL	360
Qy	300	LQIGQIRIVCLVDKSGSMATGNRLNRLNQAGQLFLQTVLQELGWSGVMTFDSAAYVOSEL	359
Db	361	KQLNSGADRLLIKHLPTVSAGGTSICSLRTAFTYIKKKYPTDGSGEIVLLTDGENTIS	420

[illegible]

```

#experimental_conditions
SUMMARY
#length 820 #molecular-weight 91464 #checksum 3104

```



```

TITLE      hypothetical protein - Synecocystis sp. (strain PCC 6803)
ORGANISM   PCC 6803
#variety
DATE       25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
          21-Aug-1998
ACCESSIONS S76691
REFERENCE   S74322
#authors   Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.;
            Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugitani, M.;
            Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
            Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo,
            S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;
            Yasuda, M.; Tabata, S.
#journal   DNA Res. (1996) 3:109-136
#title     Sequence analysis of the genome of the unicellular
            cyanobacterium Synecocystis sp. PCC6803. II. Sequence
            determination of the entire genome and assignment of
            potential protein-coding regions.
#cross-references MUID:97061201
#accession S76691
#status    preliminary
#molecule_type DNA
#residues  1-420 #label KAN
#cross-references EMBL:D64004; GB:AB001339; NID:g1001701; PID:d1011286;
            PID:g1208467
#note      the nucleotide sequence was submitted to the EMBL Data
            Library, June 1996
SUMMARY    #length 420 #molecular-weight 45849 #checksum 5250

Query Match      2.7%; Score 174; DB 2; Length 420;
Best Local Similarity 26.8%; Pred. No. 4,67e-10;
Matches          53; Conservative 53; Mismatches 80; Indels 12; Gaps 11;

Db      44  LCLVLDHSGSMD-GQPLETVKSAALGLDLRLEEDDRLSVIAFDHRKIVIVENQOVRNGA 101
Qy      307  VCLVLDKSGSMATGNRLNLRINQAGFLQTVLELGSWGVMTTFDAAHVQSELIIQNSGS 366

Db      102  AIAKATIERKAE--GGTAIDEGILKGIOEAAKGEDRVSHIFLLTDGSENEHGDRLCLK 159
Qy      367  DROTAKRLPAASAGGTSICSLRSFTVIRKYPDTGSEIVLLTDGE-DNTISG-CFN- 423

Db      160  GTVADSYKLVHTPLGFGDHWNDVLEAIAASQAGSLSYIENPSEALHTFROLFORMSNVG 219
Qy      424  EVKQSG-AI-IHTVALGPSAAQE-LEELSKMTGGTQYASDQVQN-NGLIDAFGALSSGN 479

Db      220  LTNAHLLELAPQAHIAI 237
Qy      480  GAVSQRSIQLESKG-LTL 496

RESULT      6
ENTRY       S73402 #type complete
TITLE       1-phosphofructokinase (EC 2.7.1.56) fruk - Mycoplasma
            pneumoniae (ATCC 29342) (SGC3)
ALTERNATE_NAMES
ORGANISM    hypothetical protein R02_Orf300
#variety    #formal_name Mycoplasma pneumoniae
            ATCC 29342
DATE        26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change
          17-Jul-1998
ACCESSIONS S73402
REFERENCE   S73327
#authors   Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li,
            B. C.; Herrmann, R.
#journal   Nucleic Acids Res. (1996) 24:4420-4449
#title     Complete sequence analysis of the genome of the bacterium
            Mycoplasma pneumoniae.
#cross-references MUID:97105885
#accession S73402
#status    preliminary; nucleic acid sequence not shown;
            translation not shown
#molecule_type DNA
#residues  1-300 #label HIM
#cross-references EMBL:AE000010; GB:U00089; NID:g1673729; PID:g1673730

```

```

#note      the nucleotide sequence was submitted to the EMBL Data
           Library, November 1996

GENETICS
#gene      fruk3
#genetic_code SGC3
KEYWORDS
phosphotransferase
SUMMARY
#length 300 #molecular-weight 33588 #checksum 3477

Query Match      1.8%; Score 113; DB 2; Length 300;
Best Local Similarity 26.0%; Pred. No. 2.93e-01;
Matches 19; Conservative 25; Mismatches 27; Indels 2; Gaps 2

Db 163 LLOLLNQRP-WLKKPNRDEFVAMNNSNVDOOE-LVOLLIKQFOTONTLLMSDGAOGA 220
Qy 12 ILHLLEGALSLSLLNNNGYEGIVADPNVPDETLLIOQIKDMVTQASLYLFEATGKR 71
Db 221 YFFDQOQLLFMEA 233
Qy 72 YFERNVAILIPET 84

RESULT 7
ENTRY
TITLE
ORGANISM
DATE
17-Jul-1998

ACCESSIONS
REFERENCE
#authors
Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher
C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Bar-
III, C.E.; Tekai, F.; Badcock, K.; Basham, D.; Brown, D.;
Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellon, S.; Squares, S.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrett, B.G.
Nature (1998) 393:537-544
Deciphering the biology of Mycobacterium tuberculosis from
the complete genome sequence.
#cross-references MUID:98295987
#accession
#status
preliminary; nucleic acid sequence not shown;
translation not shown

#molecule_type DNA
#residues 1-464 #label COL
#cross-references GB:AJ021897; GB:AL123456; NID:g3256022; PID:el25195
PID:g2896714
#experimental_source strain H37Rv

GENETICS
#gene
SUMMARY
#length 464 #molecular-weight 48635 #checksum 7315

Query Match      1.8%; Score 112; DB 2; Length 464;
Best Local Similarity 26.2%; Pred. No. 3.91e-01;
Matches 43; Conservative 46; Mismatches 60; Indels 15; Gaps 13

Db 217 AGRPLYVEGVGEDFWPA--AYDPS--VPDEIIAV-SDSDFDTRRLAREAMLVGGSCG 271
Qy 329 AGQFLFLQTVELGSGWGVMTFDSAAHVQSELIQNSGSDRDTLAKRLP-AAASGGTSTCS 387
Db 272 MAVVAALKVAEEAGP-DALIVVLLPDGGRGYMSKIFNDAMWSSYGFRLS-RLDGSTEOST 329
Qy 388 -GLSAFTVIRKKYPTDGSSEIVLLTDGEDNTISCGFNEVKQSG-AIHTVALGPSAAQE- 444
Db 330 GVDVLRKSGALPALVTH-PSETVRDAIGILRE-YG-VSQMPV 370
Qy 445 LEE-LSKRMGTQTYASDVQVQNNGLIDAFGALSGNGAVSQRSI 487

```



D	B	84	SENLIWIPSAIDLTFLPROLHSARS 108               :::     QY 357 SE-LIQNSGSDRDTLAKRUPAAS 380
RESULT 12			
ENTRY	B70965	#type complete	
TITLE	probable integrase - Mycobacterium tuberculosis (strain H37RV)		
ORGANISM	#formal_name Mycobacterium tuberculosis		
DATE	17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998		
ACCESSIONS	B70965		
REFERENCE	A70500		
#authors	Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigmler, K.; Gas, S.; Barry III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltham, N.; Foster, L.; Fraser, A.; Freeman, B.A.; Gair, M.J.; Galloway, P.; Hamblin, N.; Haydock, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, R.; Seeger, K.; Skelton, S.; Squares, S.; Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.		
#journal	Nature (1998) 393:537-544		
#title	Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.		
#cross-references	MUFD:98295987		
#accession	B70965		
#status	preliminary; nucleic acid sequence not shown;		
#molecule_type	DNA		
#residues	1-332 #label COL		
#cross-references	GB:Z80225; GB:AL123456; NID:g3242265; PID:e266398;		
#experimental_source	PID:g1550687		
#genome	Rv2646		
SUMMARY	#length 332 #molecular-weight 36590 #checksum 6133		
Query Match	1.7%; Score 106; DB 2; Length 332;		
Best Local Similarity	31.0%; Pred. No. 2.10e+00;		
Matches	22; Conservative 15; Mismatches 31; Indels 3; Gaps 2;		
D	B	232	RIVPISDYLAALIBDTHGVLPNGTGHLTAE-HVGKLV--SRALPGDATMHTLRHYA 288   : :: : :                 : : :       : QY 730 RTSSGGSFVADVENAPDPFPPGTDLKAELHGSLNLTWPAGDDYDHGAHKYI 789   : :: : :                 : : :       :
D	B	289	TRAYRGSHNR 299   : :: : :
Q	Y	790	IRISTILDUR 800   : :: : :
RESULT 13			
ENTRY	S30306	#type complete	
TITLE	hypothetical protein 5b - Methanobacterium thermoformicum plasmid pFVI		
ORGANISM	#formal_name Methanobacterium thermoformicum		
DATE	19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Aug-1997		
ACCESSIONS	S30306		
REFERENCE	S30302		
#authors	Noelling, J.; van Eeden, F.J.M.; Eggen, R.I.L.; de Vos, W.M.		
#journal	Nucleic Acids Res. (1992) 20:6501-6507		
#title	Modular organization of related Archaeal plasmids encoding different restriction-modification systems in Methanobacterium thermoformicum.		
#accession	S30306		
#status	preliminary; nucleic acid sequence not shown;		
#molecule_type	DNA		
#residues	1-451 ##label NOE		
#cross-references EMBL:X68366; NID:g44632; PID:g809722			
<hr/>			
#cross-references EMBL:X68366			
#note the nucleotide sequence was submitted to the EMBL Data Library, September 1992			
GENETICS			
#genome plasmid pFVI			
SUMMARY			
#length 451 #molecular-weight 49638 #checksum 5243			
Query Match 1.7%; Score 108; DB 2; Length 451;			
Best Local Similarity 25.0%; Pred. No. 1.21e+00;			
Matches 18; Conservative 23; Mismatches 30; Indels 1; Gaps 1;			
D	B	149	AGVFDDITRE-WIFETPLAGEEAATFLDCLVNRTGELSIVSVREVDEGDVVNNMAHA 207 :   :: : :   :   :   :   :   :   :   :   :   :   :   :   :   : QY 419 SGCFNEVKQSGLIHVTVALGPSAAQELELSKMTGTLQTYSOVQNGLIDAFGALS 478 :   :: : :   :   :   :   :   :   :   :   :   :   :   :   :   :
D	B	208	SVAVKTDDL 219 :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY 479 NGAVSQRSIQLE 490
RESULT 14			
ENTRY	S26441	#type complete	
TITLE	hypothetical protein 5b - Methanobacterium thermoformicum plasmid pFVI		
ORGANISM	#formal_name Methanobacterium thermoformicum		
DATE	17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Sep-1997		
ACCESSIONS	S26441		
REFERENCE	S26437		
#authors	Noelling, J.; van Eeden, F.J.M.; Eggen, R.I.L.; de Vos, W.M.		
#submission	submitted to the EMBL Data Library, September 1992		
#description	Module-like organization of related archaeal plasmids encoding different restriction-modification systems in Methanobacterium thermoformicum.		
#accession	S26441		
#molecule_type	DNA		
#residues	1-451 ##label NOE		
#cross-references	EMBL:X68366; NID:g44632; PID:g809722		
GENETICS			
#genome plasmid			
#start_codon GTG			
SUMMARY			
#length 451 #molecular-weight 49670 #checksum 5234			
Query Match 1.7%; Score 108; DB 2; Length 451;			
Best Local Similarity 25.0%; Pred. No. 1.21e+00;			
Matches 18; Conservative 23; Mismatches 30; Indels 1; Gaps 1;			
D	B	149	AGVFDDITRE-WIFETPLAGEEAATFLDCLVNRTGELSIVSVREVDEGDVVNNMAHA 207 :   :: : :   :   :   :   :   :   :   :   :   :   :   :   :   : QY 419 SGCFNEVKQSGLIHVTVALGPSAAQELELSKMTGTLQTYSOVQNGLIDAFGALS 478 :   :: : :   :   :   :   :   :   :   :   :   :   :   :   :   :
D	B	208	SVAVKTDDL 219 :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   : QY 479 NGAVSQRSIQLE 490
RESULT 15			
ENTRY	S17821	#type complete	
TITLE	protoporphyrin IX magnesium chelatase (EC 4.99.1.-) bchD - Rhodospirillum rubrum strain ATCC 35061		
ALTERNATE_NAMES	Mg chelatase 60K chain		
ORGANISM	#formal_name Rhodobacter capsulatus		
DATE	30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998		
ACCESSIONS	S17821		
REFERENCE	S17803		
#authors	Burke, D.H.; Alberti, M.; Armstrong, G.A.; Hearst, J.E.		
#journal	Submitted to the EMBL Data Library, November 1991		
#title	The complete nucleotide sequence of the 46 kb photosynthesis gene cluster of Rhodobacter capsulatus.		
#accession	S17821		
#molecule_type	DNA		
#residues	1-561 ##label EMB		



```

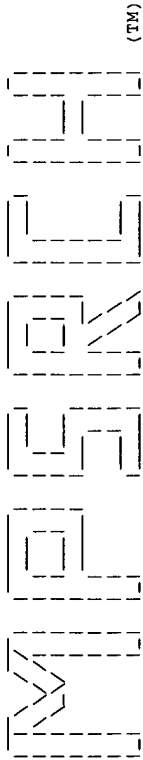
#cross-references EMBL:Z11165; NID:g46097; PID:g46120
GENETICS
#gene
KEYWORDS
SUMMARY
    bchD
    lyase
    #length 561 #molecular-weight 59740 #checksum 7380
    Query Match 1.7%; Score 111; DB 2; Length 561;
    Best Local Similarity 29.7%; Pred. No. 5.20e-01;
    Matches 35; Conservative 34; Mismatches 42; Indels 7; Gaps 7;
    Db 374 EMSDRVLIIFAVDASGSAVA-RLSEAKGAVELLGLRAYAARDHVSLITFRGTA-AQV-LL 430
    QY 301 QIGQIVCLVLDKSGSMATCNRLNRLNQAGQLFLQTVELGWSWGMVTFDSAAHVQSELI 360
    Db 431 Q-PSRSLTOT-KFOLOGLPGGGTPLASGMEMAMVTAKQARSRGMTPTIALLTDGRN 486
    QY 361 QINGSDRDTLAKRLPA-AASGGTSCSLRSFTVIRK-KYPTDGEIVLLTDGEDN 416

```

Search completed: Mon Aug 23 13:00:24 1999  
 Job time : 88 secs.

This Page Blank (uspto)

\*\*\*\*\*



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 12:56:33 1999; MasPar time 28.36 Seconds  
Tabular output not generated. 911.163 Million cell updates/sec

Title: >US-09-049-696-41  
Description: (1-914) from US09049696.pap  
Perfect Score: 6397  
Sequence: 1 MGPFKSSVFILHLLEGAL.....GIHLKIMWKWIGELQLSIA 914

Scoring table: PAM 150  
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 54.266; Variance 88.351; scale 0.614

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	3306	51.7	ECCL_BOVIN	EPITHELIAL CHLORIDE CH	0.00e+00
2	174	2.7	Y103_SYNY3	HYPOTHETICAL 45.8 KD P	2.02e-12
3	133	2.1	BCHD_RHOSH	MAGNESIUM-CHELATASE 60	3.38e-05
4	113	1.8	FRUK_MYCPN	PUTATIVE 1-PHOSPHOPRUC	4.22e-02
5	114	1.8	CH62_STRAL	60 KD CHAPERONIN 2 (PR	3.01e-02
6	118	1.8	YEBK_ECOLI	HYPOTHETICAL 63.6 KD P	7.67e-03
7	116	1.8	CBPA_CLOCL	CELLULOSE BINDING PROT	1.53e-02
8	108	1.7	YFVE_METTF	HYPOTHETICAL 49.6 KD P	2.19e-01
9	111	1.7	BCHD_RHOCA	MAGNESIUM-CHELATASE 60	8.21e-02
10	109	1.7	YDHA_ECOLI	HYPOTHETICAL PROTEIN N	1.58e-01
11	103	1.6	FLIH_ECOLI	FLAGELLAR ASSEMBLY PRO	1.06e+00
12	101	1.6	FLIH_SALTY	FLAGELLAR ASSEMBLY PRO	1.96e+00
13	102	1.6	IF6_ARATH	EUKARYOTIC TRANSLATION	1.45e+00
14	101	1.6	Y446_METJA	HYPOTHETICAL PROTEIN M	1.96e+00
15	104	1.6	IF4A_LEIBR	PROBABLE EUKARYOTIC IN	7.79e-01
16	100	1.6	YHV1_YEAST	HYPOTHETICAL 59.8 KD P	2.65e+00
17	100	1.6	PILJ_PSEAE	PILJ PROTEIN	2.65e+00
18	105	1.6	PFL_LACLC	FORMATE ACETYLTRANSFER	5.69e-01
19	104	1.6	PFL_LACLA	FORMATE ACETYLTRANSFER	7.79e-01
20	103	1.6	DPOL_HPBHE	DNA POLYMERASE (EC 2.7	1.06e+00
21	102	1.6	AKH_HAEIN	ASPARTOKINASE (EC 2.7	1.45e+00
22	100	1.6	CONT_HUMAN	CONTACTIN PRECURSOR (G	2.65e+00
23	100	1.6	CONT_MOUSE	CONTACTIN PRECURSOR (N	2.65e+00

24	104	1.6	1143	1	RPOB_PORPU	DNA-DIRECTED RNA POLYM	7.79e-01
25	102	1.6	1239	1	DPOG_HUMAN	DNA POLYMERASE GAMMA (	1.45e+00
26	97	1.5	332	1	FLGD_SALTY	BASAL-BODY ROD MODIFIC	6.45e+00
27	99	1.5	263	1	ACON_STRMU	ACONITATE HYDRATASE (E	3.58e+00
28	97	1.5	313	1	YFM4_MICOL	HYPOTHETICAL PROTEIN I	6.45e+00
29	96	1.5	328	1	Y36K_HALSP	HYPOTHETICAL 36 KD PRO	8.63e+00
30	99	1.5	333	1	F16P_RHOSH	FRUCTOSE-1,6-BISPHOSPH	3.58e+00
31	98	1.5	353	1	VM17_BORHE	VARIABLE MAJOR OUTER M	4.81e+00
32	96	1.5	395	1	VPE1_BPP2	MAJOR TAIL SHEATH PROT	8.63e+00
33	98	1.5	424	1	NABE_MYLE	PROBABLE L-ASPARTATE O	4.81e+00
34	96	1.5	468	1	ATPE_ENTHR	ATP SYNTHASE BETA CHAI	8.63e+00
35	98	1.5	499	1	PITB_ECOLI	PROBABLE LOW-AFFINITY	4.81e+00
36	97	1.5	546	1	CH60_RHILV	60 KD CHAPERONIN (PROT	6.45e+00
37	98	1.5	633	1	BEM4_YEAST	BEM4 PROTEIN.	4.81e+00
38	97	1.5	646	1	YD49_SCHPO	HYPOTHETICAL 76.2 KD P	6.45e+00
39	97	1.5	774	1	VP4_ROTHT	OUTER CAPSID PROTEIN V	4.81e+00
40	98	1.5	776	1	VP4_ROTHT	OUTER CAPSID PROTEIN V	4.81e+00
41	98	1.5	859	1	LOX1_ARATH	LIPIDYLTRANSFERASE 1 (EC	1.1
42	98	1.5	885	1	ITH3_HUMAN	INTER-ALPHA-TRYPSIN IN	4.81e+00
43	99	1.5	985	1	INVA_YERPS	INVASIN.	3.58e+00
44	96	1.5	1388	1	CAIE_HUMAN	COLLAGEN ALPHA 1(XV) C	8.63e+00
45	96	1.5	1645	1	168K_RICTY	168 KD SURFACE-LAYER P	8.63e+00

ALIGNMENTS

RESULT 1 STANDARD; PRT; 903 AA.  
ID ECCL\_BOVIN  
AC P54281;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE EPITHELIAL CHLORIDE CHANNEL PROTEIN (CALCIUM-ACTIVATED CHLORIDE CHANNEL).  
OS BOS TAURUS (BOVINE).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIOIDEA; BOVINA; BOS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=TRACHEA;  
RX MEDLINE; 96125078.  
RA CUNNINGHAM S.A., AWAYDA M.S., BUBIEN J.K., ISMAILOV I.I.,  
RA ARRATTE M.P., BERDIEV B.K., BENOS D.J., FULLER C.M.;  
RT "Cloning of an epithelial chloride channel from bovine trachea";  
RL J. BIOL. CHEM. 270:31016-31026(1995).  
CC -!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE  
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;  
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND  
CC TRANSEPITHELIAL TRANSPORT.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- TISSUE SPECIFICITY: TRACHEA.  
CC -!- PTM: PHOSPHORYLATED BY CAM-KINASE II.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
-----  
EMBL: U36445; G1184066;  
IONIC CHANNEL; TRANSMEMBRANE; ION TRANSPORT; VOLTAGE-GATED CHANNEL;  
CALCIUM CHANNEL; PHOSPHORYLATION; GLYCOPROTEIN.  
FT TRANSMEM 7 27  
FT TRANSMEN 331 351  
FT TRANSMEN 617 637  
FT TRANSMEN 853 903  
FT CARBOHYD 75 75  
FT CARBOHYD 278 278  
FT CARBOHYD 360 360  
FT CARBOHYD 372 372  
FT CARBOHYD 504 504

```

RESULT 2
ID Y103_SXNY3 STANDARD; PRT; 420 AA.
AC Q55874;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 45.8 KD PROTEIN SL0103.
GN SL0103.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96127529.
RA KANEKI T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.;
RT "sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA RES. 2:153-166(1995).
CC -I- SIMILARITY: TO E.COLI YF8K.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; D64004; G1208467; -
CC PFAM; PF00092; vwa; 1.
CC HYPOTHETICAL PROTEIN.
CC
CC SEQUENCE 420 AA; 45849 MW; 3A9D50B0 CRC32;
-----
Query Match 2.78; Score 174; DB 1; Length 420;
Best Local Similarity 26.88; Pred. No. 2.02e-12;
Matches 53; Conservative 53; Mismatches 80; Indels 12; Gaps 11;

Db 44 LCLVDHSGMD-GOPLETYKSA-LGLIDRLREDDRLSVIAFDHRAKIVINOOVRNGA 101
QY 307 VCLVLKSGSMATGNLRNLQNAQGLFLQTLVELGSGWGWTFDAAHVOSLIQINSGS 366
Db 102 AIAKAIERLKAEE--GGTAIDEGKLGIQEAAGKEDRVSHIFLLTDCGENEHDNDRCLKL 159
QY 367 DROTAKRLPAAASGGTSCSLRSATFVIRKRYPTDGSIEVLLTQGE-DNTISG-CFN- 423
Db 160 GTVASYDKLTVHVLGFGDHNQDVLEIAASAGSLSYIENPEALHTFQLFORMSNVG 219
QY 424 EVKQSG-AI-IHTVAGPSAAQE--LEELSRMTGLQTYASDVQVN-NGLIDAFALSSGN 479
Db 220 LTNALHLLLELAPOAHIAI 237
QY 480 GAVSQRSIQLESKG-LTL 496

RESULT 3
ID BCHD_RHOSH STANDARD; PRT; 558 AA.
AC Q34845;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE MAGNESIUM-CHLATASE 60 KD SUBUNIT (MG-PROTOPORPHYRIN IX CHELATASE)
DE (MG-CHELATASE SUBUNIT D).
GN BCHD.
OS RHODOBACTER SPHAEROIDES (RHODOSEUDOMONAS SPHAEROIDES).
OC BACTERIA; PROTEOBACTERIA; ALPHA SUBDIVISION; RHODOBACTER GROUP;
OC RHODOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-NCIB 8253;
RA HANSSON M., KANNANGARA C.G.;
RT "Rhodobacter Sphaeroides bchi and bchd encoding two subunits of
RT magnesium chelatase."

```

[illegible]

```
KW CHAPERONE; ATP-BINDING; MULTIGENE FAMILY.
FT INIT_MET 0
SQ SEQUENCE 539 AA; 56586 MW; 6D94EA2F CRC32;

Query Match 1.8%; Score 114; DB 1; Length 539;
Best Local Similarity 17.9%; Pred. No. 3.01e-02;
Matches 31; Conservative 54; Mismatches 81; Indels 7; Gaps 6;

Db 344 AQIRAEIENSDDYREKQERLANAGGVAVIRKAGATEVELKERKHRIEDAVRNKAA 403
Qy 353 AHVQSELIQINSGRDRLAKRLFAAASGGTSGICLSAFTVIRKYPDTGSEIVLITD 412
Db 404 VEEGIVAGGVALLQASSVEKLEEGDEATGAARAVKLALAPLKOIAVNGLGSEGVVE 463
Qy 413 GEDNTISGCFNEVQSGAILIHTVALGPSAAQLEELS-KMTGGGL-QIYASQVQNNGLID 470
Db 464 KVRNLVGHGLNATQGVYDIAEGI-LDPAK-VTRSAI-ONAAASIAALFUTT 513
Qy 471 AFGALSSGNG--AVSQRISQLESKGLTIONSQWNGTVIVDSTVGKDTLFLUIT 521

RESULT 6
ID YFEBK_ECOLI STANDARD; PRT; 575 AA.
AC P76481;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 63.6 KD PROTEIN IN ELAD-NUON INTERGENIC REGION.
GN YFEBK.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE: 97426617.
RA BLATTNER F.R., PLUNKETT G. III, BLOCH C.A., PERNA N.T., BURLAND V.,
RA RILEY M., COLLADO-VIDES J., GLASNER F.D., RODE C.K., MAYHEW G.F.,
RA GREGOR J., DAVIS N.W., KIRKPATRICK H.A., GOEDEN M.A., ROSE D.J.,
RA MAU B., SHAO Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL SCIENCE 277:1453-1474(1997).
CC -1- SIMILARITY: TO SYNECHOCYSTIS PCC 6803 SLL0103.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000317; G1788606; -.
DR ECOGENE: EG14095; YFEBK.
DR PFAM: PF00092; vwa; 1.
RW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 575 AA; 63634 MW; CBD4ED43 CRC32;

Query Match 1.8%; Score 118; DB 1; Length 575;
Best Local Similarity 27.8%; Pred. No. 7.67e-03;
Matches 30; Conservative 31; Mismatches 41; Indels 6; Gaps 6;

Db 219 FLIDTSGSMISDERL-PLIQSLKLLVKEQDNIAIVTVAGDSRI-A-IPSI-SGSHK 274
Qy 309 LVLDKSGSMATGNRLNRLNQAGQLFLQTVELGSGWGVMTFDSAHVQSELIQINSQSDR 368
Db 275 AEINAAIDSLDAESTNGGAGLELAYQOATGFKKGINRILLATDGD 322
Qy 369 DTLAKRLPA-AASGGTSGICLSAFTVIRKYPDTGSEIVLL-TDGE 414

RESULT 7
ID CBPA_CLOCL STANDARD; PRT; 1848 AA.
AC P38058;
DT 01-OCT-1994 (REL. 30, CREATED)
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CELLULOSE BINDING PROTEIN A PRECURSOR.
GN CBPA.
OS CLOSTRIDIUM CELLULOVORANS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
OC CLOSTRIDIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 9228810.
RA SHOSEVOV O., TAKAGI M., GOLDSTEIN M.A., DOI R.H.;
RT "Primary sequence analysis of Clostridium cellulovorans cellulose
RT binding protein A.";
RL PROC. NATL. ACAD. SCI. U.S.A. 89:3483-3487(1992).
CC -1- FUNCTION: BINDS TO CELLULOSE FIBERS AND COORDINATES CELLULOSE
CC ENZYMES.
CC -1- PTM: THE N-TERMINAL IS BLOCKED.
CC -1- PTM: CONTAINS ABOUT 10% OF CARBOHYDRATES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M73817; G144749; -.
DR PIR: A44140; A44140.
DR PFAM: PF00942; CBD_3; 1.
DR PFAM: PF00963; Cohesin; 9.
DR HSSP: Q06851; INBC.
KW CELLULOSE DEGRADATION; CELL WALL; REPEAT; SIGNAL; GLYCOPROTEIN.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 1848 CELLULOSE BINDING PROTEIN A.
FT DOMAIN 29 189 CELLULOSE-BINDING (BY SIMILARITY).
SQ SEQUENCE 1848 AA; 189152 MW; A10D4DDA CRC32;

Query Match 1.8%; Score 116; DB 1; Length 1848;
Best Local Similarity 24.8%; Pred. No. 1.53e-02;
Matches 26; Conservative 27; Mismatches 48; Indels 4; Gaps 4;

Db 423 PGVITNPIDNRQISVGTATVAGEITAAPVLTSTVPSGTAEQVSDATLLEVAS 482
Qy 556 PGIKAVGTWKYISLQASSQTLLTAVTSRASNA-TLPPITVTSKTNKDT-S-KPPSPLVVYAN 613
Db 483 VTAGDI-VLNPVTN-FSYTVNGVNIKLLFDDTLGSLQISKDGVF 525
Qy 614 IROGASPILRASVTALIESVNGKVTIVLELLDNGAGADATKDDGYV 658

RESULT 8
ID YPVB_METTF STANDARD; PRT; 451 AA.
AC P29578;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 49.6 KD PROTEIN (ORF5B).
OS METHANOBACTERIUM THERMOFORMICICUM.
OG PLASMID PFV1.
OC ACETABAE; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;
OC METHANOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 3848 / THF;
RX MEDLINE: 93126090.
RA NOELLING J., VAN EDEEN F.J.M., EGGEN R.I.L., DE VOS W.M.;
RT "Modular organization of related Archaeal plasmids encoding different
RT restriction-modification systems in Methanobacterium
RT thermoformicicum.";
RL NUCLEIC ACIDS RES. 20:6501-6507(1992).
```

```

QY 301 QIGRIVCLVLDKSGSWATGNRLNRLNQAGOLFLLQTVELSGSWGMVTFDSAAHVQSELI 360
Db 431 Q-PRSLSYQT-KROLQGLPGGGTPLASGMEMAMVYAKQARSRGMTPTIALLTDGRGN 486
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 361 QINSGSDRDTLAKRLPA-AASGGTSGICSLRSFTVIRK-KYPTDGSSEIVLLTDGEDN 416

RESULT 10
ID YDBA_ECOLI STANDARD; PRT; 1540 AA.
AC P33666;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN NEAR THE REPLICATION TERMINUS (FRAGMENT).
GN YDBA.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
CN ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE; 92190338.
RA MOSZER I., GLASER P., DANCHIN A.;
FT "Multiple IS insertion sequences near the replication terminus in
RT Escherichia coli K-12.";
RL BIOCHIMIE 73:1361-1374(1991).
CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC BETWEEN AMINO ACIDS 376 AND 377.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR ENBL; X62680; -: NOT ANNOTATED_CDS.
DR ECOGENE; EG11307; YDBA.
KW HYPOTHETICAL PROTEIN.
FT NON_TER 1
SQ SEQUENCE 1540 AA; 158300 MW; 1298336CB CRC32;

Query Match 1.7%; Score 109; DB 1; Length 1540;
Best Local Similarity 26.9%; Pred No. 1.58e-01;
Matches 28; Conservative 35; Mismatches 35; Indels 6; Gaps 6;

Db 412 SPIVLLATNGSATNAGTITGKVTQHSVFNKYSTGTGTSNFINNDVSSITGLVAQSNS 471
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 405 SEIVLL-TDGEDNTISGCFNE-VKQSGAIHTVALGPSAAQELE-ELSKMTGGL-QTVAS 460
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 472 LIINTDSGIIDYGRGSGVGMIAIDSTAENQCK-ITL-DSMWYDA 513
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 461 DQVQNGLLIDAFGALSSGNGAVSORISQLESKGLTLQNSQWMNG 504

RESULT 11
ID FLIH_ECOLI STANDARD; PRT; 228 AA.
AC P31068; P76325;
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FLAGELLAR ASSEMBLY PROTEIN FLIH.
GN FLIH OR FLA BIII OR FLA AII.3.
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
CN ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE; 97426617.

```

[illegible]



```

DR ENBL: AC003000. G2642164. -
SQ INITIATION FACTOR; PROTEIN BIOSYNTHESIS.
KW SEQUENCE 247 AA; 26511 MW; 37492764 CRC32;

Query Match 1.68; Score 102; DB 1; Length 247;
Best Local Similarity 28.08; Pred. No. 1.45e+00;
Matches 26; Conservative 21; Mismatches 41; Indels 5; Gaps 5;

Db 116 VALVHPKLEKDEIIISDVILGVVEY-RQTIANRELVGSYCSLSN-RGGMVHSTNTNVEEMV 173
QY 435 VALG-PSAAQEELELSKMTGLOQTAYSDQVONNGLIDAFGLSSGNGAVSORSIQLESK- 492
DB 174 ELANLVOPVLVAGTVNRCGSOIVSAGLTIVNDWTA 206
QY 493 GLT-LQNSOMNGTIVIVDSVGVKDTLFLITWTT 524

RESULT 14
ID Y446_METJA STANDARD; PRT; 361 AA.
AC Q57888;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DI HYPOTHETICAL PROTEIN MJ0446.
GN MJ0446.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
CN METHANOCOCCUS.
RC [1]
RP SEQUENCE FROM N.A. 2661 / ATCC 43067;
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RC MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.D., GEOGHANEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTERBACK T.D., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.R., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOSESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL SCIENCE 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO SYNECHOCYSTIS PCC 6803 SLL1285.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL: U67496; G1591150; -
CC TIGR: MJ0446; -
CC KW HYPOTHETICAL PROTEIN.
CC SEQUENCE 361 AA; 41984 MW; E33D396D CRC32;

Query Match 1.68; Score 101; DB 1; Length 361;
Best Local Similarity 28.68; Pred. No. 1.9e+00;
Matches 30; Conservative 24; Mismatches 41; Indels 10; Gaps 8;

Db 232 EVIIONFRACKGIPMENKPEPSP-IKMLKV-IILAKLILDDI-SIOIPNPLNR-ETGOL 286
QY 274 EVTRDSEDFKTKPTMT-QPNTFTSLQIGQIVCLVLDKSGSMATGNPLNRLNQAQGL 332
DB 287 FLLAGVD--DWGVSPLTRDYVNPAAEPWIKELRWETELGKL 329
QY 333 FLIQTVELGSWGVMTFDSA-AHVQSELIQNSDR-DTLAKRL 375

RESULT 15
ID IF4A_LEIBR STANDARD; PRT; 403 AA.

```

[illegible]

Search completed: Mon Aug 23 12:57:08 1999  
Job time : 35 secs.

**THIS PAGE BLANK (uspto)**

\*\*\*\*\*

W P E F L (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 12:57:28 1999; MasPar time 59.01 Seconds  
845.381 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-049-696-41  
Description: (1-914) from US09049696.pep  
Perfect Score: 6397  
Sequence: 1 MGPFKSSVFILHLLEGAL.....GHIILKIMWKIGELQLSIA 914

Scoring table: PAM 150  
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: sprembl9  
1:sp.archaea 2:sp.bacteria 3:sp.fungi 4:sp.human  
5:sp.invertebrate 6:sp.mammal 7:sp.mhc 8:sp.organelle  
9:sp.phage 10:sp.plant 11:sp.rodent 12:sp.unclassified  
13:sp.vertebrate 14:sp.virus

Statistics: Mean 52.622; Variance 86.917; scale 0.605

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	5088	79.5	913	11	GOB-5 PROTEIN.	0.00e+00
2	3221	50.4	901	11	CHLORIDE CHANNEL CACC.	0.00e+00
3	3206	50.1	905	6	LU-ECAM-1.	0.00e+00
4	2915	45.6	820	6	LU-ECAM-1.	0.00e+00
5	2911	45.5	794	6	LU-ECAM-1.	0.00e+00
6	1343	21.0	342	6	LU-ECAM-1.	1.78e-292
7	385	6.0	109	6	UNKNOWN PROTEIN (FRAGM	2.76e-57
8	124	1.9	97	14	NSS PROTEIN.	1.11e-03
9	119	1.9	643	5	HEAT SHOCK TRANSCRIPTI	6.60e-03
10	119	1.9	646	5	HEAT SHOCK TRANSCRIPTI	6.60e-03
11	119	1.9	658	5	HEAT SHOCK TRANSCRIPTI	6.60e-03
12	116	1.8	97	14	NON-STRUCTURAL PROTEIN	1.88e-02
13	115	1.8	97	14	NSS PROTEIN.	2.65e-02
14	113	1.8	277	14	H14-E.	5.25e-02
15	112	1.8	464	2	CYSTATHIONINE BETA-SYN	7.36e-02
16	112	1.8	1014	5	TO4A11.3 PROTEIN.	7.36e-02
17	108	1.7	97	14	NSS PROTEIN.	2.78e-01
18	107	1.7	97	14	NSS PROTEIN.	3.85e-01
19	108	1.7	263	2	HYPOTHETICAL 28.9 KD P	2.78e-01
20	110	1.7	276	14	41KBP FRAGMENT FROM LE	1.44e-01

21	106	1.7	332	2	P71945	HYPOTHETICAL 36.6 KD P	5.32e-01
22	110	1.7	629	2	O05809	HYPOTHETICAL 67.0 KD P	1.44e-01
23	106	1.7	1010	2	O45340	BRKA.	5.32e-01
24	109	1.7	1041	5	Q18031	NAD(P) TRANSYDROGENAS	2.00e-01
25	108	1.7	1149	11	P70704	CHROMAFFIN GRANULE ATP	2.78e-01
26	109	1.7	1167	2	P76859	SIMILAR TO.	2.00e-01
27	111	1.7	1280	13	O90933	NEURON-GLIA CELL ADHES	1.03e-01
28	103	1.6	268	2	O52838	PUTATIVE CARBON MONOXI	1.38e+00
29	104	1.6	288	14	O09501	32 KDA PROTEIN.	1.01e+00
30	105	1.6	290	1	O28134	CONSERVED HYPOTHETICAL	7.33e-01
31	102	1.6	293	14	O84691	MOVEMENT PROTEIN.	1.89e+00
32	102	1.6	389	2	O59232	ENDO-BETA-1,4-GLUCANAS	1.89e+00
33	102	1.6	400	2	O85465	ALKALINE CELLULASE CEL	1.89e+00
34	104	1.6	403	5	O62591	EIF-4A PROTEIN.	1.01e+00
35	105	1.6	411	2	O05011	NITROGEN FIXATION PROT	7.33e-01
36	103	1.6	422	4	O15189	EP3-VI PROTEIN.	1.38e+00
37	101	1.6	495	2	O46252	MEMBRANE-SPANNING TRAN	2.58e+00
38	104	1.6	539	2	P97086	60 KD CHAPERONIN (PROT	1.01e+00
39	101	1.6	550	4	O43405	COCH-5B2.	2.58e+00
40	104	1.6	570	10	O23212	SPLICING FACTOR HOMOLO	1.01e+00
41	104	1.6	663	2	O67708	REPAIR EXCISION NUCLEA	1.01e+00
42	103	1.6	759	2	O67406	CELLULOSE SYNTHASE CAT	1.38e+00
43	102	1.6	815	2	O47732	ORF16.	1.89e+00
44	101	1.6	1149	6	O29449	CHROMAFFIN GRANULE ATP	2.58e+00
45	103	1.6	1310	14	O41976	TEGUMENT PROTEIN/FGARA	1.38e+00

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	913 AA.
AC	O88826;			
DT	01-NOV-1998	(TREMBLREL. 08, CREATED)		
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)		
DE	GOB-5 PROTEIN.			
GN	GOB-5.			
OS	MUS MUSCULUS (MOUSE).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;			
OC	SCIUROGNATHI; MORIDAE; MURINAE; MUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=INTESTINE;			
RA	KOMIYA T., TANIGAWA Y., HIROHASHI S.;			
RT	"Cloning of a gene, gob-5, which is expressed in intestinal goblet cells in mice.;"			
RL	SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; AB017156; D1034712; "			
SQ	SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;			

Query Match	79.5%;	Score 5088;	DB 11;	Length 913;
Best Local Similarity	75.8%;	Pred. No. 0.00e+00;		
Matches	694;	Conservative 133;	Mismatches 82;	Indels 7; Gaps 5;
Db	1	MESLKSPEVILLHLLEGVLSSELIQLNNGYEGIVTAIDHVDPEDEALQHKDMVTA	60	
Qy	1	MGPFKSSVFILHLLEGALSLSLIQLNNGYEGIVTAIDNPVEDETLILQIKDMVTA	60	
Db	61	SPYLFEATGKRFYKNAIIPESWAKPEYTRPKLETFKNADVLVSTTSPGLNDEPYTE	120	
Qy	61	SLYLFEATGKRFYKNAIIPETWTKADYVRPKLETYNADVLVAESTPPGNDPEYTE	120	
Db	121	HIGACGKGIHILTPDLFAGKLLTOYQPDORTFVHEWAHFRGVFNENDEKFLSKG	180	
Qy	121	QMGNGCGKGIHILTPDFIAGKLLAEYGPQGRAVFEWAHLRWGVFEYNDEKFLYLSG	180	
Db	181	KPAVRCSAITKQVRCOGGSCITNGKVIDRVTGLYKDCNCFVDPDHPONEKASIMF	240	
Qy	181	RIQAVRCAGITNVYKCKGQSCYTK-RCTFNKVTLGYEKGEFVLQSRQTEKASIMF	239	
Db	241	NONINSVEFCTEKNHNOEAPNQNORCNLRSTWEVTOESEDKOTTPMTAQPPAPFSL	300	
Qy	240	AQHVDSEVFCTQNHKNEAPNQNORCNLRSTWEVIRDSDFKTKTPMTQPPNPFSL	299	



[illegible]



```
DR EMBL: F15082; G972064; -.
FT NON_TER 1 109
SQ SEQUENCE 109 AA; 12048 MW; AF422625 CRC32;

Query Match
Best Local Similarity 6.0%; Score 385; DB 6; Length 109;
Matches 55; Conservative 10; Mismatches 16; Indels 1; Gaps 1;

Db 1 RIVCLVLDKSGMTVGGRLKLNAGKLFLLQTVGQAWGVAFDSSAAYKSELVQINS 60
QY 305 RIVCLVLDKSGMATGNRLNQLAGQLFLQTVELGSGWGVTFDSSAAYVQSELQINS 364
Db 61 AADXRQCLPAAAYPOQPRXEPSI 82
QY 365 GSD-RDILAKELPAAASGGTSI 385

RESULT 8
ID Q90104 PRELIMINARY; PRT; 97 AA.
AC Q90104;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE HEAT SHOCK TRANSCRIPTION FACTOR.
OS SCHISTOSOMA MANSONI (BLOOD FLUKE).
OC EUKARYOTA; METAZOA; PLATHELMINTHES; TREMATODA; DIGenea; STRIGEIDIDA;
RN AM. J. TROP. MED. HYG. 54:563-569(1996).
RP SEQUENCE FROM N.A.
RC STRAIN=DAV-457;
RA BOWEN M.D.;
RA JACKSON C.F., BOWEN M.D., CHILES R.E., HARDY J.L., ELDRIDGE B.F.,
RT "Geographic distribution and serologic and genomic characterization
of Morro Bay virus, a newly recognized bunyavirus.";
RL AM. J. TROP. MED. HYG. 54:563-569(1996).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DAV-457;
RA BOWEN M.D.;
RA THESIS (1993), SCHOOL OF PUBLIC HEALTH,
RL UNIVERSITY OF CALIFORNIA AT BERKELEY.
DR EMBL: U31989; G1399280; -.
DR PFAM: PF01104; Bunya_NS-S; 1.
SQ SEQUENCE 97 AA; 10723 MW; 422A9F70 CRC32;

Query Match
Best Local Similarity 1.9%; Score 124; DB 14; Length 97;
Matches 24; Conservative 21; Mismatches 28; Indels 4; Gaps 4;

Db 20 SVLWNGNRLTLQLGSSSSPQKPRLLSRVSRQKGL-ILNLAS-GRWRLSTIIIFQQTGTI 77
QY 298 SLLOIGORIVCLVDKSGMATGNRL-NRLNQAGQLFLQTVELGSGW-VGMVTFDSSAHV 355
Db 78 QLVTTILPSTASODTLP 94
QY 356 QSELIQINSGRDRTLA 372

RESULT 9
ID Q44943 PRELIMINARY; PRT; 643 AA.
AC Q44943;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DE HEAT SHOCK TRANSCRIPTION FACTOR.
OS SCHISTOSOMA MANSONI (BLOOD FLUKE).
OC EUKARYOTA; METAZOA; PLATHELMINTHES; TREMATODA; DIGenea; STRIGEIDIDA;
RN SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.
RP SEQUENCE FROM N.A.
RC STRAIN=PUERTO RICAN;
QY 356 QSELIQINSGRDRTLA 372

Query Match
Best Local Similarity 1.9%; Score 119; DB 5; Length 646;
Matches 35; Conservative 49; Mismatches 61; Indels 12; Gaps 12;

Db 414 SPGLSDVIPSRLDGVTDSEIC-DLLNGD-SQTEQEVGNQNP-FSE-STLSKYKKEPIRK 469
QY 619 SPILRASVTALIESVNGKVTLELLDNGAGADATKDDGVYSRYFTTYDTNGRYSVK-VRA 677
Db 470 QHKIONPRCIAPFAESDPQVSTTLLQIDSLPWEKYQDSGFDFNLQVQNTENSRS 529
QY 678 LGVNAARRRVP-QQSGALYIPGWIEIEN-DEIQWNP-PRPEINKD-DVQHKQVCFSTRSS 733
Db 530 NGLIVGNEIIPVEP-ANLNACDDVQIFLRSEIPGKKL 565
QY 734 GGSFVADSV-PNAPIPLDFPPGQITD-LKAEIHGGS 768
```

```
RA LANTNER F., ZIV E., RAM D., SCHECHTER I.;
RL EUR. J. BIOCHEM. 0:0-0(1998).
DR EMBL: AF043421; G2895599; -.
DR EMBL: AF043419; G2895599; JOINED.
DR EMBL: AF043420; G2895599; JOINED.
DR PROSITE: PS00434; HSF_DOMAIN; 1.
KW HEAT SHOCK.
SQ SEQUENCE 643 AA; 72033 MW; 0706F081 CRC32;

Query Match
Best Local Similarity 1.9%; Score 119; DB 5; Length 643;
Matches 35; Conservative 49; Mismatches 61; Indels 12; Gaps 12;

Db 411 SPGLSDVIPSRLDGVTDSEIC-DLLNGD-SQTEQEVGNQNP-FSE-STLSKYKKEPIRK 466
QY 619 SPILRASVTALIESVNGKVTLELLDNGAGADATKDDGVYSRYFTTYDTNGRYSVK-VRA 677
Db 467 QHKIONPRCIAPFAESDPQVSTTLLQIDSLPWEKYQDSGFDFNLQVQNTENSRS 526
QY 678 LGVNAARRRVP-QQSGALYIPGWIEIEN-DEIQWNP-PRPEINKD-DVQHKQVCFSTRSS 733
Db 527 NGLIVGNEIIPVEP-ANLNACDDVQIFLRSEIPGKKL 562
QY 734 GGSFVADSV-PNAPIPLDFPPGQITD-LKAEIHGGS 768

RESULT 10
ID Q44944 PRELIMINARY; PRT; 646 AA.
AC Q44944;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HEAT SHOCK TRANSCRIPTION FACTOR.
OS SCHISTOSOMA MANSONI (BLOOD FLUKE).
OC EUKARYOTA; METAZOA; PLATHELMINTHES; TREMATODA; DIGenea; STRIGEIDIDA;
RN SCHISTOSOMATOIDEA; SCHISTOSOMATIDAE; SCHISTOSOMA.
RP SEQUENCE FROM N.A.
RC STRAIN=PUERTO RICAN;
RA LANTNER F., ZIV E., RAM D., SCHECHTER I.;
RL EUR. J. BIOCHEM. 0:0-0(1998).
RN [1]
RP SEQUENCE OF 16-646 FROM N.A.
RC STRAIN=PUERTO RICAN;
RX MEDLINE: 98316860.
RA LANTNER F., ZIV E., RAM D., SCHECHTER I.;
RT "Different forms of the mRNA encoding the heat-shock transcription
factor are expressed during the life cycle of the parasitic helminth
Schistosoma mansoni.";
RL EUR. J. BIOCHEM. 253:390-398(1998).
DR EMBL: AF043421; G2895600; -.
DR EMBL: AF043419; G2895600; JOINED.
DR EMBL: AF043420; G2895600; JOINED.
DR EMBL: AF043417; G2854021; -.
DR PROSITE: PS00434; HSF_DOMAIN; 1.
KW HEAT SHOCK.
SQ SEQUENCE 646 AA; 72319 MW; 272747A9 CRC32;

Query Match
Best Local Similarity 1.9%; Score 119; DB 5; Length 646;
Matches 35; Conservative 49; Mismatches 61; Indels 12; Gaps 12;

Db 414 SPGLSDVIPSRLDGVTDSEIC-DLLNGD-SQTEQEVGNQNP-FSE-STLSKYKKEPIRK 469
QY 619 SPILRASVTALIESVNGKVTLELLDNGAGADATKDDGVYSRYFTTYDTNGRYSVK-VRA 677
Db 470 QHKIONPRCIAPFAESDPQVSTTLLQIDSLPWEKYQDSGFDFNLQVQNTENSRS 529
QY 678 LGVNAARRRVP-QQSGALYIPGWIEIEN-DEIQWNP-PRPEINKD-DVQHKQVCFSTRSS 733
Db 530 NGLIVGNEIIPVEP-ANLNACDDVQIFLRSEIPGKKL 565
QY 734 GGSFVADSV-PNAPIPLDFPPGQITD-LKAEIHGGS 768
```

[illegible]



Search completed: Mon Aug 23 12:58:35 1999  
Job time : 67 secs.

**This Page Blank (uspto)**

\*\*\*\*\*

WIREH

(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:40:16 1999; MasPar time 9.59 Seconds  
42.147 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-42  
Description: (1-19) from US09049696.pep  
Perfect Score: 19  
Sequence: 1 YNDEKFYLSNGRIQAVRC 19

Scoring table: TABLE unitprotable

Gap 60

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 1.830; Variance 0.560; scale 3.267

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	6	31.6	1451	5	CCVinsavc spike prote	1.62e+01
2	5	26.3	11	22	Presenilin-1-1 residu	1.81e+02
3	5	26.3	14	27	Human cytochrome P450	1.81e+02
4	5	26.3	14	11	Human tumour cell sol	1.81e+02
5	5	26.3	14	1	Human tumour cell epi	1.81e+02
6	5	26.3	20	28	Peptide derived from	1.81e+02
7	5	26.3	25	11	Human tumour cell sol	1.81e+02
8	5	26.3	25	1	Human tumour cell epi	1.81e+02
9	5	26.3	26	11	Human tumour cell sol	1.81e+02
10	5	26.3	29	30	Monoclonal antibody N	1.81e+02
11	5	26.3	29	1	Human tumour cell epi	1.81e+02
12	5	26.3	34	22	Presenilin-1-1 residu	1.81e+02
13	5	26.3	40	4	Antigen tc-35c.	1.81e+02
14	5	26.3	47	30	Presenilin-1 fragment	1.81e+02
15	5	26.3	60	31	Measles virus haemagg	1.81e+02
16	5	26.3	79	1	Protein increasing pu	1.81e+02

17	5	26.3	91	5	R25187	21B4 gene clone produ	1.81e+02
18	5	26.3	136	35	W76228	Bacterial periplasmic	1.81e+02
19	5	26.3	142	30	W46784	Yeast NC2-alpha/DRAP1	1.81e+02
20	5	26.3	199	28	W34131	Streptococcus pneumon	1.81e+02
21	5	26.3	208	36	W71749	Human ubiquitin conj	1.81e+02
22	5	26.3	212	36	W80715	S. pneumoniae protein	1.81e+02
23	5	26.3	227	25	W35097	E. coli serine acetyl	1.81e+02
24	5	26.3	239	25	W35096	E. coli serine acetyl	1.81e+02
25	5	26.3	244	26	R64177	Carbapenem R (carR).	1.81e+02
26	5	26.3	245	25	W35095	E. coli serine acetyl	1.81e+02
27	5	26.3	248	25	W35090	E. coli serine acetyl	1.81e+02
28	5	26.3	249	25	W35089	E. coli serine acetyl	1.81e+02
29	5	26.3	250	25	W35088	E. coli serine acetyl	1.81e+02
30	5	26.3	250	25	W35080	E. coli serine acetyl	1.81e+02
31	5	26.3	255	25	W35079	E. coli serine acetyl	1.81e+02
32	5	26.3	255	25	W35087	E. coli serine acetyl	1.81e+02
33	5	26.3	256	25	W35086	E. coli serine acetyl	1.81e+02
34	5	26.3	257	25	W35085	E. coli serine acetyl	1.81e+02
35	5	26.3	258	25	W35084	E. coli serine acetyl	1.81e+02
36	5	26.3	259	25	W35083	E. coli serine acetyl	1.81e+02
37	5	26.3	263	25	W35094	E. coli serine acetyl	1.81e+02
38	5	26.3	268	25	W35093	E. coli serine acetyl	1.81e+02
39	5	26.3	270	25	W35092	E. coli serine acetyl	1.81e+02
40	5	26.3	271	25	W35091	E. coli serine acetyl	1.81e+02
41	5	26.3	273	25	W35072	E. coli serine acetyl	1.81e+02
42	5	26.3	273	25	W35070	E. coli serine acetyl	1.81e+02
43	5	26.3	273	25	W35082	E. coli serine acetyl	1.81e+02
44	5	26.3	273	25	W16618	Wild type Escherichia	1.81e+02
45	5	26.3	273	25	W35068	E. coli serine acetyl	1.81e+02
46	5	26.3	273	25	W35074	E. coli serine acetyl	1.81e+02
47	5	26.3	273	7	R37580	E. coli serine transac	1.81e+02
48	5	26.3	273	25	W35081	E. coli serine acetyl	1.81e+02
49	5	26.3	273	25	W35071	E. coli serine acetyl	1.81e+02
50	5	26.3	273	25	W35075	E. coli serine acetyl	1.81e+02
51	5	26.3	273	25	W35073	E. coli serine acetyl	1.81e+02
52	5	26.3	273	25	W35077	E. coli serine acetyl	1.81e+02
53	5	26.3	273	25	W35078	E. coli serine acetyl	1.81e+02
54	5	26.3	273	25	W35076	E. coli serine acetyl	1.81e+02
55	5	26.3	273	25	W35069	E. coli serine acetyl	1.81e+02
56	5	26.3	295	8	R41347	Nasturtium xyloglucan	1.81e+02
57	5	26.3	300	34	W77626	Staphylococcus aureus	1.81e+02
58	5	26.3	301	19	W01049	Product of 200 gene d	1.81e+02
59	5	26.3	319	8	R44770	Sequence of lipolytic	1.81e+02
60	5	26.3	335	7	R34539	M protein of attenuat	1.81e+02
61	5	26.3	357	32	W60227	Bacillus thuringiens	1.81e+02
62	5	26.3	382	20	W06517	Flavobacterium lycop	1.81e+02
63	5	26.3	383	35	W80730	Z. mobilis alcohol de	1.81e+02
64	5	26.3	383	5	R28947	Alcohol dehydrogenase	1.81e+02
65	5	26.3	383	22	W07698	Zymomonas mobilis alc	1.81e+02
66	5	26.3	387	2	R23652	Tak short fibre prote	1.81e+02
67	5	26.3	387	3	R12576	Ad41 long fibre prote	1.81e+02
68	5	26.3	400	2	R24247	Mutant human D3 dopa	1.81e+02
69	5	26.3	400	2	R24246	Human D3 dopaminergic	1.81e+02
70	5	26.3	406	13	R70981	E. coli acetic acid a	1.81e+02
71	5	26.3	422	25	W26528	Branched chain keto	1.81e+02
72	5	26.3	430	22	W05761	Presenilin-1-2 D253A	1.81e+02
73	5	26.3	431	24	W15772	Protein cognate of pr	1.81e+02
74	5	26.3	434	22	W05760	Presenilin-1-1 D257A	1.81e+02
75	5	26.3	438	22	W05756	Presenilin-1-1 delta2	1.81e+02
76	5	26.3	441	22	W20402	H. pylori cytoplasmic	1.81e+02
77	5	26.3	442	22	W20624	H. pylori cytoplasmic	1.81e+02
78	5	26.3	446	3	R14551	D-3 dopaminergic rece	1.81e+02
79	5	26.3	446	3	R15349	D-3 dopaminergic rece	1.81e+02
80	5	26.3	445	8	R41727	Peptide derived from	1.81e+02
81	5	26.3	453	30	W23965	Human presenilin-1.	1.81e+02
82	5	26.3	463	29	W42375	Human presenilin I ge	1.81e+02
83	5	26.3	463	22	W05734	Presenilin-1-2	1.81e+02
84	5	26.3	463	27	W22948	Presenilin-1 VRSQ var	1.81e+02
85	5	26.3	463	21	W12376	Human S182 gene produ	1.81e+02
86	5	26.3	463	21	W11840	Early onset Alzheimer	1.81e+02
87	5	26.3	467	22	W05750	Presenilin-1-1 C263R	1.81e+02
88	5	26.3	467	30	W23964	Human presenilin-1	1.81e+02
89	5	26.3	467	22	W05751	Presenilin-1-1 P264L	1.81e+02

90	5	26.3	457 22	W05754	Presenilin-1-1 A285V	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
91	5	26.3	457 22	W05736	Presenilin-1-1 A79X m	1.81e+02	1.81e+02	Human NMDA R1e recept	1.81e+02
92	5	26.3	457 22	W05739	Presenilin-1-1 Y115H	1.81e+02	1.81e+02	Human excitatory amin	1.81e+02
93	5	26.3	457 22	W05748	Presenilin-1-1 A246E	1.81e+02	1.81e+02	NMDAR1-1.	1.81e+02
94	5	26.3	457 22	W05738	Presenilin-1-1 V96F m	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
95	5	26.3	457 22	W05758	Presenilin-1-1 L392V	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
96	5	26.3	457 22	W05759	Presenilin-1-1 C410Y	1.81e+02	1.81e+02	C-terminal amidation	1.81e+02
97	5	26.3	457 22	W05773	Presenilin-1-1.	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
98	5	26.3	457 24	W21777	Human mutant S182 gen	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
99	5	26.3	457 22	W05737	Presenilin-1-1 V82L m	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
100	5	26.3	457 29	W41430	PS1/467 protein.	1.81e+02	1.81e+02	NMDA receptor channel	1.81e+02
101	5	26.3	457 22	W05740	Presenilin-1-1 M139T/	1.81e+02	1.81e+02	NMDA receptor channel	1.81e+02
102	5	26.3	457 22	W05749	Presenilin-1-1 A260V	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
103	5	26.3	457 22	W05752	Presenilin-1-1 P267S	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
104	5	26.3	467 21	W11839	Human early onset Alz	1.81e+02	1.81e+02	Human NMDA R1d recept	1.81e+02
105	5	26.3	467 24	W21716	Human S182 gene, PS1	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
106	5	26.3	467 22	W05741	Presenilin-1-1 I143T	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
107	5	26.3	467 22	W05743	Presenilin-1-1 H183R/	1.81e+02	1.81e+02	N-methyl-D-aspartate	1.81e+02
108	5	26.3	467 22	W05744	Presenilin-1-1 L171P	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
109	5	26.3	467 22	W05757	Presenilin-1-1 G384A	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
110	5	26.3	467 22	W05755	Presenilin-1-1 L286V	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
111	5	26.3	467 22	W05747	Presenilin-1-1 I231T	1.81e+02	1.81e+02	Human NMDA R1a recept	1.81e+02
112	5	26.3	467 33	W56770	Homo sapiens PS-1	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
113	5	26.3	467 22	W05745	Presenilin-1-1 G209V	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
114	5	26.3	467 22	W05746	Presenilin-1-1 I211T	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
115	5	26.3	467 22	W05742	Presenilin-1-1 M146L/	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
116	5	26.3	467 22	W05753	Presenilin-1-1 E280A/	1.81e+02	1.81e+02	Protein encoded by se	1.81e+02
117	5	26.3	471 21	W20184	H. pylori cytoplasmic	1.81e+02	1.81e+02	Type B alpha-amidation	1.81e+02
118	5	26.3	486 34	W69945	GlutRAGln amidotrans	1.81e+02	1.81e+02	C-terminal amidation	1.81e+02
119	5	26.3	493 13	R73262	Human cytochrome P450	1.81e+02	1.81e+02	Porcine PAM-8.	1.81e+02
120	5	26.3	493 18	R93169	Human cytochrome P450	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
121	5	26.3	493 17	R81467	Human derived cytochr	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
122	5	26.3	510 12	R60863	Hydrogen peroxide-gen	1.81e+02	1.81e+02	C-terminal amidation	1.81e+02
123	5	26.3	520 23	W13666	Fragment of dbpA gene	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
124	5	26.3	550 10	R53079	Acetyl coenzyme A. ch	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
125	5	26.3	550 28	W38416	Human acyl-coenzyme A	1.81e+02	1.81e+02	Bovine peptidyl-glyci	1.81e+02
126	5	26.3	550 29	W43409	Human acylcoenzyme A.	1.81e+02	1.81e+02	Equine C-terminal ami	1.81e+02
127	5	26.3	555 35	W49098	Amino acid sequence o	1.81e+02	1.81e+02	Human N-methyl-D-aspa	1.81e+02
128	5	26.3	562 23	W12096	Adenovirus E1A associ	1.81e+02	1.81e+02	Human NMDA receptor s	1.81e+02
129	5	26.3	564 6	R30613	Babesia bovis immunor	1.81e+02	1.81e+02	C-terminal amidation	1.81e+02
130	5	26.3	565 18	R97981	Babesia microzoite sur	1.81e+02	1.81e+02	Glucocorticoid prote	1.81e+02
131	5	26.3	617 8	R42391	Chicago 2 haemaggluti	1.81e+02	1.81e+02	Rat receptor type-pro	1.81e+02
132	5	26.3	621 35	W76213	Human ELL protein.	1.81e+02	1.81e+02	Human Down syndrome-c	1.81e+02
133	5	26.3	622 12	R67424	Alginic acid lyase.	1.81e+02	1.81e+02	Human PTP-OB.	1.81e+02
134	5	26.3	627 9	R49590	GAT-B transporter.	1.81e+02	1.81e+02	Human protein tyrosin	1.81e+02
135	5	26.3	627 8	R41229	GAT-3 GABA transporte	1.81e+02	1.81e+02	Mus musculus factor V	1.81e+02
136	5	26.3	632 8	R41232	GAT-3 transporter.	1.81e+02	1.81e+02	Murine factor VIII.	1.81e+02
137	5	26.3	640 35	W76212	Human ELL2 protein.	1.81e+02	1.81e+02	C terminal peptide of	1.72e+03
138	5	26.3	650 8	R41730	Peptide derived from	1.81e+02	1.81e+02	Human tumour necrosis	1.72e+03
139	5	26.3	693 1	P94856	Expression plasmid pU	1.81e+02	1.81e+02	Breast cancer-associ	1.72e+03
140	5	26.3	715 15	R80482	Recombinant bovine ad	1.81e+02	1.81e+02	Consensus sequence of	1.72e+03
141	5	26.3	734 23	W13667	Streptomyces viridosp	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
142	5	26.3	776 2	R06400	Lambda gt10ch2 encode	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
143	5	26.3	777 38	R85578	Human N-methyl-D-aspa	1.81e+02	1.81e+02	Protein sequence enco	1.72e+03
144	5	26.3	777 12	R66044	Human N-methyl-D-aspa	1.81e+02	1.81e+02	Human TNF-alpha conve	1.72e+03
145	5	26.3	780 21	W11819	Candida albicans topo	1.81e+02	1.81e+02	Leuko T-2.	1.72e+03
146	5	26.3	793 4	R20743	Murine receptor-type	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
147	5	26.3	793 18	W02282	Murine receptor type	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
148	5	26.3	802 4	R20744	Human receptor-type p	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
149	5	26.3	821 23	W13668	DhpA-mel chimeric pro	1.81e+02	1.81e+02	Substrate for TNFalph	1.72e+03
150	5	26.3	823 23	W13668	Candida albicans topo	1.81e+02	1.81e+02	Sequence of TNF conve	1.72e+03
151	5	26.3	853 2	R03222	Type A alpha-amidatin	1.81e+02	1.81e+02	Peptide component of	1.72e+03
152	5	26.3	853 2	R06426	Protein encoded by se	1.81e+02	1.81e+02	Human hepatitis B spe	1.72e+03
153	5	26.3	853 2	R11110	Rat C-terminal amidat	1.81e+02	1.81e+02	Peptide containing th	1.72e+03
154	5	26.3	854 38	W85586	Human N-methyl-D-aspa	1.81e+02	1.81e+02	CDR2 of the heavy cha	1.72e+03
155	5	26.3	854 38	W85586	Human N-methyl-D-aspa	1.81e+02	1.81e+02	Interferon-receptor b	1.72e+03
156	5	26.3	866 2	R66052	Human NMDA receptor s	1.81e+02	1.81e+02	Polymerase enhancing	1.72e+03
157	5	26.3	866 2	R06379	Lambda gt10ch101 enco	1.81e+02	1.81e+02	Neuropeptide (5).	1.72e+03
158	5	26.3	870 38	W85585	Human N-methyl-D-aspa	1.81e+02	1.81e+02	N-terminal of follicl	1.72e+03
159	5	26.3	870 12	R66051	Human NMDA receptor s	1.81e+02	1.81e+02	Peptide Cp2A derive	1.72e+03
160	5	26.3	875 38	W85584	N-methyl-D-aspartate	1.81e+02	1.81e+02	Peptide C.p.2A derive	1.72e+03
161	5	26.3	875 1	P94854	C-terminal prepro-C-t	1.81e+02	1.81e+02	D. maculata antigen 5	1.72e+03
162	5	26.3	875 13	R73053	Peptidyl C-terminal a	1.81e+02	1.81e+02		

236	4	21.1	4	21.1	309	144	30	W47082	Salmonella Sef14 prot	1.72e+03	1.72e+03
237	4	21.1	4	21.1	310	147	34	W77738	Staphylococcus aureus	1.72e+03	1.72e+03
238	4	21.1	4	21.1	311	147	37	W72404	Arabidopsis thaliana	1.72e+03	1.72e+03
239	4	21.1	4	21.1	312	148	11	R58876	Rat-411 cadherin part	1.72e+03	1.72e+03
240	4	21.1	4	21.1	313	148	17	R87118	Proctadherin clone R	1.72e+03	1.72e+03
241	4	21.1	4	21.1	314	150	26	W22491	Phaffia derived glyce	1.72e+03	1.72e+03
242	4	21.1	4	21.1	315	154	3	R13496	P. denitrificans COB E	1.72e+03	1.72e+03
243	4	21.1	4	21.1	316	155	38	R68547	Gibbon interleukin-3	1.72e+03	1.72e+03
244	4	21.1	4	21.1	317	155	29	W55394	H. pylori ORF 14ee103	1.72e+03	1.72e+03
245	4	21.1	4	21.1	318	160	21	W07574	Porcine kringle 1-2	1.72e+03	1.72e+03
246	4	21.1	4	21.1	319	161	21	W14262	Mouse epimorphin trun	1.72e+03	1.72e+03
247	4	21.1	4	21.1	320	161	21	W14559	Human epimorphin trun	1.72e+03	1.72e+03
248	4	21.1	4	21.1	321	162	34	W77584	Staphylococcus aureus	1.72e+03	1.72e+03
249	4	21.1	4	21.1	322	164	38	W82399	Human G3BP protein	1.72e+03	1.72e+03
250	4	21.1	4	21.1	323	164	13	R67587	Bovine male enhanced	1.72e+03	1.72e+03
251	4	21.1	4	21.1	324	165	23	W23571	Salmonella enteritidis	1.72e+03	1.72e+03
252	4	21.1	4	21.1	325	165	12	R62752	Sefa sequence	1.72e+03	1.72e+03
253	4	21.1	4	21.1	326	166	26	P30684	Consensus human leuko	1.72e+03	1.72e+03
254	4	21.1	4	21.1	327	166	26	P30686	Consensus human leuko	1.72e+03	1.72e+03
255	4	21.1	4	21.1	328	167	2	R11533	Consensus human leuko	1.72e+03	1.72e+03
256	4	21.1	4	21.1	329	167	2	R11532	Sequence of polyepit	1.72e+03	1.72e+03
257	4	21.1	4	21.1	330	169	3	P60218	Mouse interleukin-2 (	1.72e+03	1.72e+03
258	4	21.1	4	21.1	331	169	3	P61032	Moubatin platelet agg	1.72e+03	1.72e+03
259	4	21.1	4	21.1	332	171	8	R41488	Human adult testis se	1.72e+03	1.72e+03
260	4	21.1	4	21.1	333	172	39	R82001	Murine male enhanced	1.72e+03	1.72e+03
261	4	21.1	4	21.1	334	174	13	R67586	Human tumour necrosis	1.72e+03	1.72e+03
262	4	21.1	4	21.1	335	177	2	P60417	Human SP-18 surfactan	1.72e+03	1.72e+03
263	4	21.1	4	21.1	336	181	39	W88199	Recombinant apoaequor	1.72e+03	1.72e+03
264	4	21.1	4	21.1	337	189	12	R65911	Aequorin E1285 mutant	1.72e+03	1.72e+03
265	4	21.1	4	21.1	338	189	19	W05058	Aequorin G122A mutant	1.72e+03	1.72e+03
266	4	21.1	4	21.1	339	189	19	W05059	Sequence encoded by 1	1.72e+03	1.72e+03
267	4	21.1	4	21.1	340	189	5	P20110	Human HPDv78 protein	1.72e+03	1.72e+03
268	4	21.1	4	21.1	341	191	38	W73377	Human TNF-alpha conve	1.72e+03	1.72e+03
269	4	21.1	4	21.1	342	193	39	P90067	Sequence encoding nov	1.72e+03	1.72e+03
270	4	21.1	4	21.1	343	196	3	P60273	Apoaequorin D1245 mut	1.72e+03	1.72e+03
271	4	21.1	4	21.1	344	196	19	W05054	Recombinant apoaequor	1.72e+03	1.72e+03
272	4	21.1	4	21.1	345	196	12	R65910	Wild type apoaequorin	1.72e+03	1.72e+03
273	4	21.1	4	21.1	346	196	33	W62540	H. pylori cell envelo	1.72e+03	1.72e+03
274	4	21.1	4	21.1	347	198	22	W20437	Vespid antigen 5s fro	1.72e+03	1.72e+03
275	4	21.1	4	21.1	348	199	8	R43885	Antigen 2 from cluste	1.72e+03	1.72e+03
276	4	21.1	4	21.1	349	203	28	W35689	Expressed antigen for	1.72e+03	1.72e+03
277	4	21.1	4	21.1	350	210	38	W89867	Streptococcus pneumon	1.72e+03	1.72e+03
278	4	21.1	4	21.1	351	210	38	W89890	H. pylori ORF 03gp208	1.72e+03	1.72e+03
279	4	21.1	4	21.1	352	210	33	W55080	P. sulcata type I pol	1.72e+03	1.72e+03
280	4	21.1	4	21.1	353	210	29	W55504	Human calcium channel	1.72e+03	1.72e+03
281	4	21.1	4	21.1	354	213	39	W82690	Syntaxin 1A deletion	1.72e+03	1.72e+03
282	4	21.1	4	21.1	355	214	39	W82690	Murine STIF protein	1.72e+03	1.72e+03
283	4	21.1	4	21.1	356	216	33	W63157	Human secreted protei	1.72e+03	1.72e+03
284	4	21.1	4	21.1	357	222	35	W69404	Influenza haemaggluti	1.72e+03	1.72e+03
285	4	21.1	4	21.1	358	222	18	R87536	Phenazine gene cluste	1.72e+03	1.72e+03
286	4	21.1	4	21.1	359	222	16	R77358	Pseudomonas aureofaci	1.72e+03	1.72e+03
287	4	21.1	4	21.1	360	222	15	R06436	Cysteine proteinase f	1.72e+03	1.72e+03
288	4	21.1	4	21.1	361	224	1	R06436	Haloacetate dehalogen	1.72e+03	1.72e+03
289	4	21.1	4	21.1	362	224	1	R75202	Tyrosine phosphatase	1.72e+03	1.72e+03
290	4	21.1	4	21.1	363	225	15	R75202	Protein sequence enco	1.72e+03	1.72e+03
291	4	21.1	4	21.1	364	228	19	W46879	Tazarotene induced eno	1.72e+03	1.72e+03
292	4	21.1	4	21.1	365	228	19	R98742	Human colon specific	1.72e+03	1.72e+03
293	4	21.1	4	21.1	366	228	8	W06548	Staphylococcal entero	1.72e+03	1.72e+03
294	4	21.1	4	21.1	367	228	8	R45013	Human proTNF-alpha pr	1.72e+03	1.72e+03
295	4	21.1	4	21.1	368	233	39	W90065	Human tumour necrosis	1.72e+03	1.72e+03
296	4	21.1	4	21.1	369	233	3	P60074	Sequence of tumour ne	1.72e+03	1.72e+03
297	4	21.1	4	21.1	370	233	3	P60531	Human tumour necrosis	1.72e+03	1.72e+03
298	4	21.1	4	21.1	371	233	3	P60531	Novel cancer-destruct	1.72e+03	1.72e+03
299	4	21.1	4	21.1	372	233	3	P61179	Human tumour necrosis	1.72e+03	1.72e+03
300	4	21.1	4	21.1	373	233	3	P61172	Antigen 1 from cluste	1.72e+03	1.72e+03
301	4	21.1	4	21.1	374	233	18	W00454	Protein encoded by cl	1.72e+03	1.72e+03
302	4	21.1	4	21.1	375	234	39	W89972	Human mature IGFBP-1	1.72e+03	1.72e+03
303	4	21.1	4	21.1	376	234	39	W89826	Human mature IGFBP-1	1.72e+03	1.72e+03
304	4	21.1	4	21.1	377	234	11	R62046	Sequence of tumour ne	1.72e+03	1.72e+03
305	4	21.1	4	21.1	378	235	3	P60530	Amino acid sequence o	1.72e+03	1.72e+03
306	4	21.1	4	21.1	379	235	31	W27951	Mouse tumour necrosis	1.72e+03	1.72e+03
307	4	21.1	4	21.1	380	235	16	R88369			
308	4	21.1	4	21.1	381						
309	4	21.1	4	21.1	382						
310	4	21.1	4	21.1	383						
311	4	21.1	4	21.1	384						
312	4	21.1	4	21.1	385						
313	4	21.1	4	21.1	386						
314	4	21.1	4	21.1	387						
315	4	21.1	4	21.1	388						
316	4	21.1	4	21.1	389						
317	4	21.1	4	21.1	390						
318	4	21.1	4	21.1	391						
319	4	21.1	4	21.1	392						
320	4	21.1	4	21.1	393						
321	4	21.1	4	21.1	394						
322	4	21.1	4	21.1	395						
323	4	21.1	4	21.1	396						
324	4	21.1	4	21.1	397						
325	4	21.1	4	21.1	398						
326	4	21.1	4	21.1	399						
327	4	21.1	4	21.1	400						
328	4	21.1	4	21.1	401						
329	4	21.1	4	21.1	402						
330	4	21.1	4	21.1	403						
331	4	21.1	4	21.1	404						
332	4	21.1	4	21.1	405						
333	4	21.1	4	21.1	406						
334	4	21.1	4	21.1	407						
335	4	21.1	4	21.1	408						
336	4	21.1	4	21.1	409						
337	4	21.1	4	21.1	410						
338	4	21.1	4	21.1	411						
339	4	21.1	4	21.1	412						
340	4	21.1	4	21.1	413						
341	4	21.1	4	21.1	414						
342	4	21.1	4	21.1	415						
343	4	21.1	4	21.1	416						
344	4	21.1	4	21.1	417						
345	4	21.1	4	21.1	418						
346	4	21.1	4	21.1	419						
347	4	21.1	4	21.1	420						
348	4	21.1	4	21.1	421						
349	4	21.1	4	21.1	422						
350	4	21.1	4	21.1	423						
351	4	21.1	4	21.1	424						
352	4	21.1	4	21.1	425						
353	4	21.1	4	21.1	426						
354	4	21.1	4	21.1	427						
355	4	21.1	4	21.1	428						
356	4	21.1	4	21.1	429						
357	4	21.1	4	21.1	430						
358	4	21.1	4	21.1	431						
359	4	21.1	4	21.1	432						
360	4	21.1	4	21.1	433						
361	4	21.1	4	21.1	434						
362	4	21.1	4	21.1							

382	4	21.1	240	8	R39697	Myasthenic antigenic	1.72e+03	455	4	21.1	326	8	R41541	BL5R product.	1.72e+03
383	4	21.1	240	14	R76698	ShET1 enterotoxin enc	1.72e+03	456	4	21.1	326	17	R85860	WD-40 domain-contg. b	1.72e+03
384	4	21.1	241	31	W53948	Human NSP10 protein B	1.72e+03	457	4	21.1	327	3	P50046	Hybrid trpE-SS14 prot	1.72e+03
385	4	21.1	242	16	R79871	Anti-EGFR single chai	1.72e+03	458	4	21.1	331	39	W67889	Human secreted protei	1.72e+03
386	4	21.1	242	16	R79872	Anti-EGFR single chai	1.72e+03	459	4	21.1	331	33	W63752	A. nidulans IPNS prot	1.72e+03
387	4	21.1	243	4	R24133	Human HLA DC-alpha.	1.72e+03	460	4	21.1	335	18	R97628	Human SLAM1 T-cell co	1.72e+03
388	4	21.1	244	38	W83019	A murine OCIF-binding	1.72e+03	461	4	21.1	336	23	W18051	DNAX accessory molecu	1.72e+03
389	4	21.1	244	16	R79868	Anti-EGFR single chai	1.72e+03	462	4	21.1	337	35	W72034	HSV-2 strain S85 Cont	1.72e+03
390	4	21.1	246	35	W89781	Staphylococcus aureus	1.72e+03	463	4	21.1	338	29	W55522	H. pylori ORF 14ap102	1.72e+03
391	4	21.1	246	35	W64260	Human amyloid beta-pr	1.72e+03	464	4	21.1	339	26	W21771	Protein encoded by OR	1.72e+03
392	4	21.1	248	38	W73284	Human proteasome subu	1.72e+03	465	4	21.1	340	17	R85859	WD-40 domain-contg. b	1.72e+03
393	4	21.1	251	28	W41149	Maleate isomerase seq	1.72e+03	466	4	21.1	340	17	R85863	WD-40 domain-contg. b	1.72e+03
394	4	21.1	251	39	W89804	Staphylococcus aureus	1.72e+03	467	4	21.1	340	17	R85861	WD-40 domain-contg. d	1.72e+03
395	4	21.1	251	28	W27298	Human J143-1 secreted	1.72e+03	468	4	21.1	341	14	R72964	Pig kidney cell mutar	1.72e+03
396	4	21.1	252	22	W20502	H. pylori surface mem	1.72e+03	469	4	21.1	341	16	R77361	Cysteine proteinase f	1.72e+03
397	4	21.1	255	29	W35461	Mouse RNA-binding pro	1.72e+03	470	4	21.1	341	35	W69760	Acetobacter xylinum C	1.72e+03
398	4	21.1	256	39	W95463	LS170 polypeptide seq	1.72e+03	471	4	21.1	341	5	P80056	Sequence encoded by t	1.72e+03
399	4	21.1	257	19	W02162	Lactococcus lactis su	1.72e+03	472	4	21.1	342	12	R63215	Human alpha-1,3-fucos	1.72e+03
400	4	21.1	259	1	P91868	Recombinant IGF bindi	1.72e+03	473	4	21.1	343	28	W35001	Bacillus thermoloevo	1.72e+03
401	4	21.1	260	3	P61317	Serratia nuclease pro	1.72e+03	474	4	21.1	344	38	W83139	BR20 protein.	1.72e+03
402	4	21.1	261	38	W88759	Polypeptide fragment	1.72e+03	475	4	21.1	344	1	P90046	Human nonspecific cro	1.72e+03
403	4	21.1	262	26	W23073	Thermococcus esterase	1.72e+03	476	4	21.1	344	2	P94006	Sequence of CEA antiq	1.72e+03
404	4	21.1	263	28	W31584	Emeria tenella sporo	1.72e+03	477	4	21.1	344	9	R54715	Carcinoembryonic anti	1.72e+03
405	4	21.1	264	4	R20500	Cell growth inhibitor	1.72e+03	478	4	21.1	344	20	W06874	Carcinoembryonic anti	1.72e+03
406	4	21.1	264	18	R88723	LiPi1.	1.72e+03	479	4	21.1	345	9	R47929	Acetobacter acetati res	1.72e+03
407	4	21.1	266	25	W24930	Syntaxin 1A N-termin	1.72e+03	480	4	21.1	346	18	W00245	Human elav-like neuro	1.72e+03
408	4	21.1	269	2	P70758	Pig pancreas elastase	1.72e+03	481	4	21.1	350	21	W14532	Human chimeric fucosy	1.72e+03
409	4	21.1	269	31	W53322	P-selectin ligand pro	1.72e+03	482	4	21.1	350	21	W14532	Human chimeric fucosy	1.72e+03
410	4	21.1	269	19	R98459	mGA-methyltransferase	1.72e+03	483	4	21.1	353	23	W19215	Swine growth hormone	1.72e+03
411	4	21.1	269	19	W02161	Lactococcus lactis su	1.72e+03	484	4	21.1	353	11	P91374	Vaccinia WR serine pr	1.72e+03
412	4	21.1	274	21	W15426	Subtilisin Carlsberg	1.72e+03	485	4	21.1	354	11	R57013	Human A transferase (	1.72e+03
413	4	21.1	275	5	R25203	Human PH-203 protein	1.72e+03	486	4	21.1	357	15	R77394	Fragment of alkaline	1.72e+03
414	4	21.1	277	13	R66478	Human epimorphine pro	1.72e+03	487	4	21.1	358	27	W35127	R. pipiens recombinan	1.72e+03
415	4	21.1	279	7	R36554	Mouse epimorphin iso	1.72e+03	488	4	21.1	358	30	W23968	Connexin protein Cx40	1.72e+03
416	4	21.1	280	12	R63760	sacS anti-terminator.	1.72e+03	489	4	21.1	358	9	R46980	Fish growth hormone.	1.72e+03
417	4	21.1	282	39	W67850	Human secreted protei	1.72e+03	490	4	21.1	358	27	W35130	R. pipiens recombinan	1.72e+03
418	4	21.1	283	8	R44155	Moraxella bovis mboA	1.72e+03	491	4	21.1	359	9	R45933	Alpha-(1-3)Fuc-RVI.	1.72e+03
419	4	21.1	284	34	W75860	Human secretory prote	1.72e+03	492	4	21.1	359	21	W14523	Human chimeric fucosy	1.72e+03
420	4	21.1	286	38	W83136	Human CEA protein imm	1.72e+03	493	4	21.1	359	32	W60578	Human neuron-specific	1.72e+03
421	4	21.1	286	1	P82933	SAP(Phe).	1.72e+03	494	4	21.1	359	22	W20927	H. pylori surface or	1.72e+03
422	4	21.1	286	20	W06871	Carcinoembryonic anti	1.72e+03	495	4	21.1	359	22	W14522	Human chimeric fucosy	1.72e+03
423	4	21.1	287	16	R92038	Human epimorphin isoA	1.72e+03	496	4	21.1	360	21	W14522	Human chimeric fucosy	1.72e+03
424	4	21.1	288	25	W24927	Rat syntaxin 1A.	1.72e+03	497	4	21.1	360	27	W35128	R. pipiens recombinan	1.72e+03
425	4	21.1	288	29	W55289	H. pylori ORF 14cel107	1.72e+03	498	4	21.1	360	21	W14513	Human chimeric fucosy	1.72e+03
426	4	21.1	288	16	R92041	Mouse epimorphin isoA	1.72e+03	499	4	21.1	361	23	W19217	Human growth hormone	1.72e+03
427	4	21.1	288	18	R96421	Rat syntaxin 1A.	1.72e+03	500	4	21.1	361	21	W14518	Human chimeric fucosy	1.72e+03
428	4	21.1	288	35	W60874	Manic-depressive illn	1.72e+03	501	4	21.1	361	21	W14519	Human chimeric fucosy	1.72e+03
429	4	21.1	288	7	R66553	Mouse epimorphin iso	1.72e+03	502	4	21.1	362	23	W19610	Human growth hormone	1.72e+03
430	4	21.1	289	13	R66479	Wild type mouse epimo	1.72e+03	503	4	21.1	362	1	R04211	Deduced amino acid se	1.72e+03
431	4	21.1	289	28	W43421	Rat syntaxin 3 protei	1.72e+03	504	4	21.1	363	18	W69396	Pyrolinin gene pro	1.72e+03
432	4	21.1	292	39	W80988	Amino acid sequence o	1.72e+03	505	4	21.1	363	35	R87532	Pseudomonas fluoresce	1.72e+03
433	4	21.1	293	22	W20287	H. pylori inner membr	1.72e+03	506	4	21.1	364	23	W19613	Rat growth hormone se	1.72e+03
434	4	21.1	294	33	W69596	NR-KB receptor activa	1.72e+03	507	4	21.1	365	26	W23536	CDK inhibitory fusion	1.72e+03
435	4	21.1	296	36	W80622	S. pneumoniae protein	1.72e+03	508	4	21.1	368	2	P93996	Transmembrane (TM)-4	1.72e+03
436	4	21.1	298	29	W55339	H. pylori ORF 14cel109	1.72e+03	509	4	21.1	369	20	R05384	Xylene oxygenase gene	1.72e+03
437	4	21.1	299	20	W06881	Decay accelerating fa	1.72e+03	510	4	21.1	369	20	W06822	Turkey herpes virus v	1.72e+03
438	4	21.1	300	39	W73518	KSHV assembly protei	1.72e+03	511	4	21.1	371	19	W01113	RNA polymerase primar	1.72e+03
439	4	21.1	305	39	W73519	KSHV assembly protei	1.72e+03	512	4	21.1	374	9	R45939	A polysyltransferase	1.72e+03
440	4	21.1	306	35	W60673	Manic-depressive illn	1.72e+03	513	4	21.1	375	4	R22096	Actin.	1.72e+03
441	4	21.1	313	3	P60071	Sequence of the TrpE	1.72e+03	514	4	21.1	375	3	P61532	Sequence of beta-acti	1.72e+03
442	4	21.1	316	38	W83017	Osteoclastogenesis in	1.72e+03	515	4	21.1	375	8	R49248	Actin.	1.72e+03
443	4	21.1	316	38	W88749	Secreted protein enco	1.72e+03	516	4	21.1	375	4	R22026	A. chrysogenium actin.	1.72e+03
444	4	21.1	316	38	W83194	Human osteoprotegerin	1.72e+03	517	4	21.1	375	29	W44332	Class I acyl-ACP thio	1.72e+03
445	4	21.1	317	38	W88758	Polypeptide fragment	1.72e+03	518	4	21.1	376	23	W19799	Gamma-smooth muscle i	1.72e+03
446	4	21.1	317	1	R07366	Polypeptide encoded b	1.72e+03	519	4	21.1	379	27	W35126	R. pipiens recombinan	1.72e+03
447	4	21.1	317	2	R11603	Protein involved in g	1.72e+03	520	4	21.1	381	39	W35205	Decay accelerating fa	1.72e+03
448	4	21.1	318	24	W27506	Bacillus thuringiens	1.72e+03	521	4	21.1	381	1	P82934	SAP(Phe).	1.72e+03
449	4	21.1	318	39	W95502	C. acetobutylicum inc	1.72e+03	522	4	21.1	381	1	R04833	Includes a 6kD human	1.72e+03
450	4	21.1	321	4	R22045	Human arginase.	1.72e+03	523	4	21.1	381	3	R14446	Pulmonary surfactant	1.72e+03
451	4	21.1	322	1	R05306	Carcino embryonic ant	1.72e+03	524	4	21.1	381	3	R82982	Human SP18 deduced fr	1.72e+03
452	4	21.1	323	7	R37381	E.coli O-acetylserine	1.72e+03	525	4	21.1	382	35	W76190	Actinoplanes sp. acar	1.72e+03
453	4	21.1	323	29	W55360	H. pylori ORF 03gei05	1.72e+03	526	4	21.1	384	2	R24296	Regulatory protein va	1.72e+03
454	4	21.1	325	19	R99226	Recombinant human alp	1.72e+03	527	4	21.1	384	23	W14076	S.thermophilus expopol	1.72e+03

528	4	21.1	385 10	R54633	Consensus sequence of	1.72e+03	601	4	21.1	472 4	P40675	Sequence encoded by t	1.72e+03
529	4	21.1	385 18	W00081	Human integrase inter	1.72e+03	602	4	21.1	473 10	R52604	Allinase encoded by c	1.72e+03
530	4	21.1	389 38	W87550	Intein (protein intro	1.72e+03	603	4	21.1	473 39	W90062	Maize MS45 protein.	1.72e+03
531	4	21.1	389 38	W85454	An intein of Thermoco	1.72e+03	604	4	21.1	473 13	R66610	Human cdc25C gene pro	1.72e+03
532	4	21.1	389 19	R99222	Recombinant human alp	1.72e+03	605	4	21.1	473 22	W15239	Atrazine chlorohydril	1.72e+03
533	4	21.1	389 16	R79921	Simian-human hybrid T	1.72e+03	606	4	21.1	476 14	R82250	Mature human wild typ	1.72e+03
534	4	21.1	390 1	R05666	Hybrid transforming g	1.72e+03	607	4	21.1	477 2	R06457	v-PA.alphal.	1.72e+03
535	4	21.1	390 9	R46227	Human pre-TGF-beta-1.	1.72e+03	608	4	21.1	478 1	P90459	Alpha-factor profibro	1.72e+03
536	4	21.1	390 1	R05492	Chimeric simian TGF-b	1.72e+03	609	4	21.1	478 2	R40083	Profibrolase product	1.72e+03
537	4	21.1	391 16	R83054	Transforming growth f	1.72e+03	610	4	21.1	479 14	R72604	Human calcium channel	1.72e+03
538	4	21.1	391 1	R81362	Human transforming gr	1.72e+03	611	4	21.1	483 28	W31405	Bacillus amyloliquefa	1.72e+03
539	4	21.1	391 32	W37992	Mutant Aspergillus or	1.72e+03	612	4	21.1	484 39	W89274	Granulocytic Ehrlichi	1.72e+03
540	4	21.1	393 19	R99223	Recombinant human alp	1.72e+03	613	4	21.1	485 3	R15507	Tomato ACC synthase e	1.72e+03
541	4	21.1	394 2	R06486	v-PA.gamma.	1.72e+03	614	4	21.1	485 3	R15506	Tomato ACC synthase	1.72e+03
542	4	21.1	394 39	W86363	Mouse DNAX toll-like	1.72e+03	615	4	21.1	485 29	W47313	Tomato ACC synthase.	1.72e+03
543	4	21.1	394 2	R05078	Aspergillopepsin A ge	1.72e+03	616	4	21.1	485 35	W61470	Staphylococcus aureus	1.72e+03
544	4	21.1	397 38	W86031	Rhizobiaceae phenol d	1.72e+03	617	4	21.1	485 35	W61465	Staphylococcus aureus	1.72e+03
545	4	21.1	398 4	R20505	Mutant Human alpha-1-	1.72e+03	618	4	21.1	486 22	W20586	H. pylori cytoplasmic	1.72e+03
546	4	21.1	398 4	R20507	Mutant Human alpha-1-	1.72e+03	619	4	21.1	491 24	W13373	Recombinant Sp 100 an	1.72e+03
547	4	21.1	400 2	R12131	ORF 3 of IgG light ch	1.72e+03	620	4	21.1	493 7	R35230	Human ALT.	1.72e+03
548	4	21.1	402 8	R44335	Alpha-antichymotrypsi	1.72e+03	621	4	21.1	493 3	R13392	P. falciparum sporozo	1.72e+03
549	4	21.1	402 12	R67259	Alpha-1-antichymotryp	1.72e+03	622	4	21.1	495 30	W46904	A human mutant alanin	1.72e+03
550	4	21.1	405 2	R04082	Gene product of south	1.72e+03	623	4	21.1	496 10	R53114	Crucifer 1-aminocyclo	1.72e+03
551	4	21.1	405 1	P90458	Snake venom fibrolase	1.72e+03	624	4	21.1	496 18	R94605	Human TIE-2 ligand 2	1.72e+03
552	4	21.1	405 6	R28840	Hela cell fucosyltran	1.72e+03	625	4	21.1	497 7	R36740	Human tyrosine hydrox	1.72e+03
553	4	21.1	406 4	R21349	Human Tryptophan oxyg	1.72e+03	626	4	21.1	497 10	R50184	Human tyrosine hydrox	1.72e+03
554	4	21.1	409 39	W89440	Mycoplasma arginini a	1.72e+03	627	4	21.1	503 9	R45318	Ced-3 (G360S).	1.72e+03
555	4	21.1	409 39	W89441	Mycoplasma arginini a	1.72e+03	628	4	21.1	503 9	R53281	Ced-3 (G360S).	1.72e+03
556	4	21.1	410 11	R55791	Cell death hid protei	1.72e+03	629	4	21.1	503 9	R45308	Ced-3 mutant P344.	1.72e+03
557	4	21.1	413 22	W20789	H. pylori cytoplasmic	1.72e+03	630	4	21.1	503 9	R45302	Ced-3 mutant H315.	1.72e+03
558	4	21.1	415 26	W22982	Canine herpesvirus g1	1.72e+03	631	4	21.1	503 9	R45277	Ced-3 mutant M234.	1.72e+03
559	4	21.1	415 27	W34319	Beta-2 adrenergic rece	1.72e+03	632	4	21.1	503 9	R45282	Ced-3 mutant N250.	1.72e+03
560	4	21.1	416 13	R69554	Human lysosomal membr	1.72e+03	633	4	21.1	503 9	R47466	Ced-3.	1.72e+03
561	4	21.1	417 4	R22044	Carcino embryonic ant	1.72e+03	634	4	21.1	503 9	R53287	Ced-3 (E483K).	1.72e+03
562	4	21.1	419 20	W01504	Wild-type human pancr	1.72e+03	635	4	21.1	503 9	R45317	Ced-3 mutant G360.	1.72e+03
563	4	21.1	421 39	W26744	S. carnosus nitrate r	1.72e+03	636	4	21.1	503 9	R45262	Ced-3.	1.72e+03
564	4	21.1	421 39	W86326	Kidney injury associa	1.72e+03	637	4	21.1	503 9	R53286	Ced-3 (A466V).	1.72e+03
565	4	21.1	422 17	R85871	WD-40 domain-contg. O	1.72e+03	638	4	21.1	503 9	R53285	Ced-3 (A449V).	1.72e+03
566	4	21.1	423 27	W27635	Streptococcus pneumon	1.72e+03	639	4	21.1	503 19	R98758	Nematode Ced-3 cell d	1.72e+03
567	4	21.1	424 39	W86150	Human growth-related	1.72e+03	640	4	21.1	503 9	R45291	Ced-3 mutant N285.	1.72e+03
568	4	21.1	424 27	W31707	Bovine coronavirus E3	1.72e+03	641	4	21.1	503 9	R45290	Ced-3 mutant K283.	1.72e+03
569	4	21.1	428 17	R95909	Opacity associated pr	1.72e+03	642	4	21.1	503 9	R53279	Ced-3 (L27F).	1.72e+03
570	4	21.1	431 2	R06459	v-PA.beta.	1.72e+03	643	4	21.1	503 19	R98754	Nematode Ced-3 cell d	1.72e+03
571	4	21.1	433 18	W00086	Human Ras-related pro	1.72e+03	644	4	21.1	503 9	R45325	Ced-3 mutant V452.	1.72e+03
572	4	21.1	435 31	W53917	SLG-13 protein.	1.72e+03	645	4	21.1	503 9	R45320	Ced-3 mutant R429.	1.72e+03
573	4	21.1	436 4	R21547	Human Tryptophan-2,3-	1.72e+03	646	4	21.1	503 9	R45276	Ced-3 mutant S183.	1.72e+03
574	4	21.1	437 14	R76521	Alpha-amyase-8-C.	1.72e+03	647	4	21.1	503 9	R45312	Ced-3 mutant K351.	1.72e+03
575	4	21.1	438 25	W18209	Staphylococcus aureus	1.72e+03	648	4	21.1	503 9	R45313	Ced-3 mutant Q356.	1.72e+03
576	4	21.1	438 14	R76519	Alpha-amyase-6-C.	1.72e+03	649	4	21.1	503 9	R45310	Ced-3 mutant K349.	1.72e+03
577	4	21.1	440 1	P94774	Membrane bound decay	1.72e+03	650	4	21.1	503 9	R45289	Ced-3 mutant V280.	1.72e+03
578	4	21.1	440 27	W27483	Human glycoposphatid	1.72e+03	651	4	21.1	503 9	R45285	Ced-3 mutant G261.	1.72e+03
579	4	21.1	441 10	R54223	L.lactis branched ami	1.72e+03	652	4	21.1	503 9	R45295	Ced-3 mutant F298.	1.72e+03
580	4	21.1	442 3	R12331	TrpE/AR androgen-bind	1.72e+03	653	4	21.1	503 9	R45294	Ced-3 mutant M291.	1.72e+03
581	4	21.1	442 27	W27442	Oil seed rape cysteine	1.72e+03	654	4	21.1	503 9	R45287	Ced-3 mutant G277.	1.72e+03
582	4	21.1	446 12	R66448	C. jejuni hippuricase	1.72e+03	655	4	21.1	503 9	R45280	Ced-3 mutant I447.	1.72e+03
583	4	21.1	447 33	W060830	E. coli glutamate deh	1.72e+03	656	4	21.1	503 9	R45315	Ced-3 mutant C358.	1.72e+03
584	4	21.1	447 20	W08092	Glutamine dehydroge	1.72e+03	657	4	21.1	503 9	R45309	Ced-3 mutant L346.	1.72e+03
585	4	21.1	448 1	P93343	Gene encoding the 43	1.72e+03	658	4	21.1	503 9	R53280	Ced-3 (G65R).	1.72e+03
586	4	21.1	450 20	W06090	Human TATA-binding pr	1.72e+03	659	4	21.1	504 28	R35008	Vibrio harveyi endogl	1.72e+03
587	4	21.1	451 12	R63088	Breast tumor kinase.	1.72e+03	660	4	21.1	509 8	R47348	Human PH-20 (derivd)	1.72e+03
588	4	21.1	453 8	R40856	43kd regression assoc	1.72e+03	661	4	21.1	509 26	W34557	Thermococcus 9M2 glyc	1.72e+03
589	4	21.1	457 32	W59009	Leucocin A ATP access	1.72e+03	662	4	21.1	509 28	W44280	Human sperm surface p	1.72e+03
590	4	21.1	459 29	W38454	Mouse RNA-binding pro	1.72e+03	663	4	21.1	509 5	R25202	Human PH-20 protein H	1.72e+03
591	4	21.1	459 29	W38457	Human RNA-binding pro	1.72e+03	664	4	21.1	509 14	R77437	BGP (1-314)/CEA (490-	1.72e+03
592	4	21.1	463 21	W04360	Human cathepsin C hom	1.72e+03	665	4	21.1	511 14	R77438	BGP (1-314)/CEA (490-	1.72e+03
593	4	21.1	464 12	R67566	Tomato spotted wilt v	1.72e+03	666	4	21.1	512 18	R93173	Human cytochrome P450	1.72e+03
594	4	21.1	464 14	R65166	Biliary glycoprotein.	1.72e+03	667	4	21.1	512 18	R93171	Human cytochrome P450	1.72e+03
595	4	21.1	465 2	R11423	Tryptophanase.	1.72e+03	668	4	21.1	518 38	W85472	PS18 protein encoded	1.72e+03
596	4	21.1	467 30	W23966	Mouse presenilin-1 ho	1.72e+03	669	4	21.1	519 22	W20640	H. pylori transporter	1.72e+03
597	4	21.1	469 39	W89797	Staphylococcus aureus	1.72e+03	670	4	21.1	525 32	W58550	Human excitatory amin	1.72e+03
598	4	21.1	489 39	W89560	Triticum sp. cysteine	1.72e+03	671	4	21.1	526 38	W83138	FL-CEA protein.	1.72e+03
599	4	21.1	470 30	W55398	Protein SEQ ID NO:229	1.72e+03	672	4	21.1	526 38	R54714	Carcinoembryonic anti	1.72e+03
600	4	21.1	472 39	W89559	Triticum sp. cysteine	1.72e+03	673	4	21.1	527 38	W81560	Truncated RSV F prote	1.72e+03





820	4	21.1	941	2	R07478	Cellulase.	1.72e+03	893	4	21.1	1203	31	W57315	Human endothelial cit	1.72e+03
821	4	21.1	943	34	W64378	Mycobacterium tubercu	1.72e+03	894	4	21.1	1205	16	R77353	Endothelial nitrogen	1.72e+03
822	4	21.1	947	37	W81745	M. tuberculosis immun	1.72e+03	895	4	21.1	1217	10	R49832	Beat-galactosidase/he	1.72e+03
823	4	21.1	947	39	W81168	Transcriptional regul	1.72e+03	896	4	21.1	1218	25	W18354	Proliferation and dif	1.72e+03
824	4	21.1	953	2	R07167	105KD PTK protein of	1.72e+03	897	4	21.1	1225	30	W52289	Homo sapiens cdo tumo	1.72e+03
825	4	21.1	956	6	R32356	Excitatory amino acid r	1.72e+03	898	4	21.1	1229	10	R54074	CrVET5.	1.72e+03
826	4	21.1	968	23	W21721	SH2-A.	1.72e+03	899	4	21.1	1229	39	W87633	CrVET5 protein sequen	1.72e+03
827	4	21.1	970	20	R99460	Biosynthetic enzyme o	1.72e+03	900	4	21.1	1229	26	W35259	Bacillus thuringiens	1.72e+03
828	4	21.1	973	13	R75708	Eph-related PKCek10	1.72e+03	901	4	21.1	1229	22	W17699	CrVET5.	1.72e+03
829	4	21.1	984	32	W59960	Human transient recep	1.72e+03	902	4	21.1	1242	30	W52287	Rattus norvegicus cdo	1.72e+03
830	4	21.1	982	22	W20980	H. pylori secreted or	1.72e+03	903	4	21.1	1252	14	R80530	B. sphaericus SLP.	1.72e+03
831	4	21.1	982	20	W11304	Receptor-type tyrosin	1.72e+03	904	4	21.1	1257	9	R46627	Neurocan core protein	1.72e+03
832	4	21.1	988	13	R75710	Eph-related PKCek10	1.72e+03	905	4	21.1	1267	1	R04232	Rickettsia rickettsii	1.72e+03
833	4	21.1	994	24	W63365	Mouse Nuk tyrosine ki	1.72e+03	906	4	21.1	1296	8	R41198	CT.	1.72e+03
834	4	21.1	999	17	R87511	Human c-mer protoonc	1.72e+03	907	4	21.1	1312	35	W71295	Human homologue of ye	1.72e+03
835	4	21.1	999	26	W26682	Bovine lysosomal alph	1.72e+03	908	4	21.1	1334	10	R50290	lktA::lacZ fusion pro	1.72e+03
836	4	21.1	1005	32	W60181	B. thuringiensis tran	1.72e+03	909	4	21.1	1337	11	R63125	IGF-I receptor 957A v	1.72e+03
837	4	21.1	1009	5	R36206	Type B human platelet	1.72e+03	910	4	21.1	1350	17	R91296	Drosophila nitric oxi	1.72e+03
838	4	21.1	1011	13	R75709	Eph-related PKCek5+	1.72e+03	911	4	21.1	1358	36	W69494	Haemagglutinin protei	1.72e+03
839	4	21.1	1014	1	P80810	Sequence of pol prote	1.72e+03	912	4	21.1	1358	17	R96032	P. gingivalis hagD ha	1.72e+03
840	4	21.1	1022	3	P30097	Fusion protein of 59-	1.72e+03	913	4	21.1	1363	27	W51706	Bovine coronavirus E2	1.72e+03
841	4	21.1	1025	13	R70126	Serum opacity factor	1.72e+03	914	4	21.1	1367	17	R91429	Human type I insulin-	1.72e+03
842	4	21.1	1036	4	R20599	ROD HIV-2 polymerase.	1.72e+03	915	4	21.1	1367	27	W37692	Human insulin-like gr	1.72e+03
843	4	21.1	1039	38	W85014	Stat5-green fluoresce	1.72e+03	916	4	21.1	1370	3	P60005	Sequence encoded by h	1.72e+03
844	4	21.1	1045	38	W85050	Stat5-green fluoresce	1.72e+03	917	4	21.1	1389	21	W14051	TIM protein splice va	1.72e+03
845	4	21.1	1045	39	W86354	Human DNAX toll-like	1.72e+03	918	4	21.1	1394	16	R92768	Adhesion and penetra	1.72e+03
846	4	21.1	1055	23	W13055	HIV-2 provirus-encode	1.72e+03	919	4	21.1	1399	8	R38698	S-PRV-055 TGE virus g	1.72e+03
847	4	21.1	1066	17	R91300	TATA box binding prot	1.72e+03	920	4	21.1	1400	9	R44514	MLL amino acid sequen	1.72e+03
848	4	21.1	1068	9	R46294	Ptdins 3-kinase l10 k	1.72e+03	921	4	21.1	1403	23	W20033	Neuronal apoptosis in	1.72e+03
849	4	21.1	1085	39	W89563	Bovine parathyroid ca	1.72e+03	922	4	21.1	1429	8	R44489	Sequence of all or pa	1.72e+03
850	4	21.1	1085	18	R95607	RRP3 telomerase-assoc	1.72e+03	923	4	21.1	1430	31	W56786	Rat neuronal nitric o	1.72e+03
851	4	21.1	1087	17	R36028	P. gingivalis haemagg	1.72e+03	924	4	21.1	1433	16	R77382	Neuronal nitrogen mon	1.72e+03
852	4	21.1	1087	36	W59486	Haemagglutinin protel	1.72e+03	925	4	21.1	1447	26	R68553	Deleted in colorectal	1.72e+03
853	4	21.1	1088	39	W88253	Gracilariaopsis lemame	1.72e+03	926	4	21.1	1452	8	R44400	Canine coronavirus 1-	1.72e+03
854	4	21.1	1088	14	R72398	Glucan Lyase 1.	1.72e+03	927	4	21.1	1454	9	R42468	Feline enteric corona	1.72e+03
855	4	21.1	1089	2	R08267	Platelet derived grow	1.72e+03	928	4	21.1	1454	9	R42467	Feline infectious per	1.72e+03
856	4	21.1	1089	2	R06910	Alpha type PDGF recep	1.72e+03	929	4	21.1	1454	9	R42464	Feline infectious per	1.72e+03
857	4	21.1	1091	39	W88254	Gracilariaopsis lemame	1.72e+03	930	4	21.1	1454	5	R24511	Prod. of the S genome	1.72e+03
858	4	21.1	1091	14	R72399	Glucan Lyase 1.	1.72e+03	931	4	21.1	1454	9	R42475	FEV/FIPV chimeric sp	1.72e+03
859	4	21.1	1091	14	R70634	Fungus-infected G.lem	1.72e+03	932	4	21.1	1523	38	W87554	A Tfu DNA polymerase	1.72e+03
860	4	21.1	1092	39	W82571	Gracilariaopsis lemame	1.72e+03	933	4	21.1	1528	31	W57487	Murine multidrug resi	1.72e+03
861	4	21.1	1098	21	W13866	Chimeric protein #1.	1.72e+03	934	4	21.1	1531	25	W23597	Human LYST1 short iso	1.72e+03
862	4	21.1	1115	23	W22478	Duffy antigen binding	1.72e+03	935	4	21.1	1545	31	W55966	Human canalicular mul	1.72e+03
863	4	21.1	1120	39	W21642	Mouse elf protein.	1.72e+03	936	4	21.1	1577	16	R91047	Alpha-D-glucosyltrans	1.72e+03
864	4	21.1	1121	7	R36728	Ubiquitin fusion prot	1.72e+03	937	4	21.1	1588	9	R46605	Malarial PfEMP3 epito	1.72e+03
865	4	21.1	1125	30	W52288	Rattus norvegicus cdo	1.72e+03	938	4	21.1	1594	5	R81183	Sequence of the peplo	1.72e+03
866	4	21.1	1128	1	R05107	Sequence encoded by R	1.72e+03	939	4	21.1	1644	25	W25049	BRCA2 cancer suscepti	1.72e+03
867	4	21.1	1139	31	W37779	Rattus norvegicus Cla	1.72e+03	940	4	21.1	1663	16	R94029	Human modified C3 (r1	1.72e+03
868	4	21.1	1142	13	R70831	Human JAK1 kinase.	1.72e+03	941	4	21.1	1663	16	R94028	Human C3 precursor.	1.72e+03
869	4	21.1	1144	16	R77360	Inducible nitrogen mo	1.72e+03	942	4	21.1	1663	27	W34613	Human C3 protein muta	1.72e+03
870	4	21.1	1144	31	W51246	Inducible nitric oxid	1.72e+03	943	4	21.1	1663	27	W34614	Human C3 protein muta	1.72e+03
871	4	21.1	1147	18	W02571	Rat inducible nitric	1.72e+03	944	4	21.1	1663	9	R46608	Plasmodium falciparum	1.72e+03
872	4	21.1	1150	25	W22471	Streptococcal C4a pep	1.72e+03	945	4	21.1	1663	27	W34627	Human C3 protein muta	1.72e+03
873	4	21.1	1151	38	W37344	Rat alphas protein se	1.72e+03	946	4	21.1	1663	27	W34628	Human C3 protein muta	1.72e+03
874	4	21.1	1151	32	W23059	Rat beta 2 integrin a	1.72e+03	947	4	21.1	1663	27	W34616	Human C3 protein muta	1.72e+03
875	4	21.1	1151	32	W60001	Rat alpha d polypetid	1.72e+03	948	4	21.1	1663	27	W34615	Human C3 protein muta	1.72e+03
876	4	21.1	1153	16	R88464	Hepatocyte inducible	1.72e+03	949	4	21.1	1663	27	W34610	Human C3 protein muta	1.72e+03
877	4	21.1	1153	28	W36113	Human inducible nitri	1.72e+03	950	4	21.1	1667	27	W34626	Human C3 protein muta	1.72e+03
878	4	21.1	1153	10	R55764	Sequence encoded by t	1.72e+03	951	4	21.1	1704	28	W34843	Arg-gingipain high mo	1.72e+03
879	4	21.1	1161	38	W73345	Rat alphas protein se	1.72e+03	952	4	21.1	1704	13	R70188	Arg-gingipain-2 prepo	1.72e+03
880	4	21.1	1161	33	W55104	Rat beta-integrin alp	1.72e+03	953	4	21.1	1732	36	W24787	Prk antigenic protei	1.72e+03
881	4	21.1	1169	19	W06417	Antiscarab pest toxin	1.72e+03	954	4	21.1	1732	36	W69487	Haemagglutinin protel	1.72e+03
882	4	21.1	1169	18	R96126	Bacillus thuringiens	1.72e+03	955	4	21.1	1738	10	R50971	Norwalk virus strain	1.72e+03
883	4	21.1	1171	38	W85037	Green fluorescent pro	1.72e+03	956	4	21.1	1805	5	R27204	Rat nestin.	1.72e+03
884	4	21.1	1174	38	W57438	Human protein tyrosin	1.72e+03	957	4	21.1	1829	27	W29322	DNA polymerase with 3	1.72e+03
885	4	21.1	1181	38	W85038	NFAT1-green fluoresce	1.72e+03	958	4	21.1	1829	29	W38757	Phosphatidyl inositol	1.72e+03
886	4	21.1	1186	32	W51030	Mus musculus dty prot	1.72e+03	959	4	21.1	1876	26	W22607	Platenolide synthase	1.72e+03
887	4	21.1	1186	27	W31503	Nematode toxin 80JJI	1.72e+03	960	4	21.1	2035	11	R57141	Host cell factor prot	1.72e+03
888	4	21.1	1186	23	W21694	Bacillus thuringiens	1.72e+03	961	4	21.1	2104	39	W89578	Calcium permeable vol	1.72e+03
889	4	21.1	1187	25	W18352	Proliferation and dif	1.72e+03	962	4	21.1	2104	34	W57773	Musca domestica volta	1.72e+03
890	4	21.1	1189	25	W26623	Signalling inositol p	1.72e+03	963	4	21.1	2105	39	W89577	Calcium permeable vol	1.72e+03
891	4	21.1	1198	33	W64384	S. cerevisiae L9470.2	1.72e+03	964	4	21.1	2105	34	W57772	Musca domestica volta	1.72e+03
892	4	21.1	1199	5	R26188	HVTA antigen.	1.72e+03	965	4	21.1	2154	39	W61639	Mouse elf-1 protein.	1.72e+03

```
966 4 21.1 2166 33 W48714 RSV isolate 2B wild-t 1.72e+03
967 4 21.1 2166 33 W48717 RSV vaccine 2B20L L p 1.72e+03
968 4 21.1 2166 33 W48719 RSV revertant 2B20L T 1.72e+03
969 4 21.1 2227 27 W34074 Hepatitis A virus HM- 1.72e+03
970 4 21.1 2233 33 W48712 HPIV-3 FRH1 cp45 vacc 1.72e+03
971 4 21.1 2233 33 W48713 HPIV-3 Vero cp45 vacc 1.72e+03
972 4 21.1 2237 33 W63142 Human calcium channel 1.72e+03
973 4 21.1 2240 11 R67819 Acetyl CoA carboxylas 1.72e+03
974 4 21.1 2289 25 W14987 Protein derived from 1.72e+03
975 4 21.1 2322 22 W15566 Melanoma-associated c 1.72e+03
976 4 21.1 2329 25 W25038 Partial BRCA2 cancer 1.72e+03
977 4 21.1 2339 6 R33549 Sequence of the alpha 1.72e+03
978 4 21.1 2670 15 R88125 Rat IP3 receptor. 1.72e+03
979 4 21.1 2818 21 W13280 Human neurofibromin. 1.72e+03
980 4 21.1 2861 26 W27227 Human TRIO phosphopro 1.72e+03
981 4 21.1 2893 35 W71556 Helicobacter polypept 1.72e+03
982 4 21.1 2938 11 R59923 GAP protein Irai. 1.72e+03
983 4 21.1 3084 2 P94758 Sequence of mouse lam 1.72e+03
984 4 21.1 3224 31 W54235 Human Nup358 protein. 1.72e+03
985 4 21.1 3396 8 R43662 DEN1-S275/90 (ECACC V 1.72e+03
986 4 21.1 3418 26 W23287 Human breast and ovar 1.72e+03
987 4 21.1 3567 8 R44431 eryA region polypepti 1.72e+03
988 4 21.1 3639 8 R40227 ACVS. 1.72e+03
989 4 21.1 3647 1 R05041 Filamentous haemagglu 1.72e+03
990 4 21.1 3672 27 W31950 Human bg protein asso 1.72e+03
991 4 21.1 3801 27 W31949 Human bg protein asso 1.72e+03
992 4 21.1 4536 28 W41262 Apolipoprotein B-100. 1.72e+03
993 4 21.1 4544 9 R47861 Alpha 2-Macroglobulin 1.72e+03
994 4 21.1 4550 26 W23716 Platenolide synthase 1.72e+03
995 4 21.1 4550 26 W22806 Platenolide synthase 1.72e+03
996 4 21.1 4630 23 W19529 Streptomyces venezuel 1.72e+03
997 4 21.1 4655 29 W43311 Human calcium sensor 1.72e+03
998 4 21.1 4655 29 W43314 Human parathyroid cal 1.72e+03
999 4 21.1 4655 29 W43313 Human kidney calcium 1.72e+03
1000 4 21.1 5072 2 R11510 Ryanodine receptor de 1.72e+03
```

## ALIGNMENTS

```
RESULT 1
ID R27819 standard; Protein; 1451 AA.
AC R27819;
DT 15-MAR-1993 (first entry)
DE CCVinsavc spike protein.
KW Canine coronavirus; CCV; spike protein; SP; CCV-6; Insave-1; dogs;
KW Liverpool C54 strain; vaccine; cats; feline peritonitis virus.
OS Canine corona virus.
PN EP-510773-A.
PD 28-OCT-1992.
PF 22-APR-1992; 201136.
PR 25-APR-1991; EP-303737.
PA (ALKU ) AK20 NV.
PI Brown TDK, Horschburgh BC;
DR N-PSDB; Q29685.
PT New nucleic acid for canine coronavirus spike protein - derived
PT vectors and transformed cells, useful in vaccines
PS Claim 1: Page 25-31; 559p; English.
CC The sequences given in R27818-20 are polypeptide determinants of
CC a canine coronavirus (CCV) spike protein (SP). These sequences are
CC derived from either the CCV-6, Insave-1 or Liverpool C54 strains.
CC These sequences, the nucleotides encoding them or recombinant vectors
CC containing the DNA sequences can be useful in vaccines to protect dogs
CC against CCV infection. They can also be used to protect cats against
CC infection by feline peritonitis virus.
SQ Sequence 1451 AA;

Query Match 31.6%; Score 6; DB 5; Length 1451;
Best Local Similarity 100.0%; Pred. No. 1.62e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1263 dekfyl 1268
|||||

RESULT 2
ID W05770 standard; peptide; 11 AA.
AC W05770;
DT 28-JUL-1997 (first entry)
DE Presenilin-1 residues 50-60.
KW Presenilin-1; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
OS Homo sapiens.
PD W09634099-A2.
PN 31-OCT-1996.
PR 29-APR-1996; CA0263.
PR 28-APR-1995; US-431048.
PR 28-JUN-1995; US-496841.
PR 31-JUL-1995; US-509359.
PA (HSCR-) HSC RES & DEV LP.
PI (UTOR ) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
PI WPI; 96-497631/49.
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
PS Claim 71; Page -; 178pp; English.
CC W05768-W05788 represent antigenic fragments of the human presenilin-1-1
CC protein (see W05733 for wild type sequence). W05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot spot
CC regions. Mutations in PS genes are implicated in familial Alzheimer's
CC disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC schizophrenia, depression etc., so detection of mutations in the DNA
CC encoding the wild type sequences can be used for diagnosis of these
CC diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are useful
CC as models for drug screening. The antibodies can also be used e.g. for
CC affinity purification and in immunoassays.
SQ Sequence 11 AA;

Query Match 26.3%; Score 5; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 1.81e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 lsngsr 5
|||||
QY 9 LSNGR 13

RESULT 3
ID W30849 standard; peptide; 14 AA.
AC W30849;
DT 27-MAR-1998 (first entry)
DE Human cytochrome P450 peptide fragment 4.
KW CYP2E1; human cytochrome P450; antigen; antibody; diagnosis;
KW cancer; diabetes.
OS Homo sapiens.
PD J09274042-A.
PN 21-OCT-1997.
PR 05-APR-1996; 134493.
PR 05-APR-1996; JP-134493.
PA (NCHK ) NICHIREI KK.
DR WPI; 98-005435/01.
PT An antibody against human CYP2E1 - useful for the diagnosis and
PT treatment of cancer and diabetes
PS Claim 5; Page 2; 6pp; Japanese.
CC Peptides w30846-53 are fragments of CYP2E1, a human cytochrome P450.
CC These peptides were used as antigens to raise antibodies which recognise
CC CYP2E1. The antibodies are used for identifying CYP2E1 in a sample. This
CC is useful for the diagnosis and treatment of cancer and diabetes.
CC
```

SQ Sequence 14 AA;  
Query Match 26.3%; Score 5; DB 27; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 6 ndekf 10  
| | | | |  
QY 3 NDEKF 7  
RESULT 4  
ID R57740 standard; peptide: 14 AA.  
AC R57740;  
DT 21-MAR-1995 (first entry)  
DE Human tumour cell sol. protein antigen partial sequence.  
KW Tumour; cancer; soluble protein; antigen; antibody; Ab;  
KW monoclonal antibody; mAb; epitope; mAb 16-88;  
KW colon carcinoma cell line.  
OS Homo sapiens.  
PN US5338832-A.  
PD 16-AUG-1994.  
PF 02-JUL-1987; 069478.  
PR 02-JUL-1987; US-069478.  
PR 28-FEB-1989; US-343475.  
PR 13-AUG-1992; US-929842.  
PA (ALKU) AKZO NV.  
PI Berman ML, Bos ES, Hanna MG, Haspel MV, Pomato N;  
DR WPI: 94-263329/32.  
PT Pure human tumour cell soluble protein antigen - comprises human  
PT tumour cell epitope to which monoclonal antibody 16-88 binds  
PS Claim 4; Col 12; 14pp; English.  
CC The human tumour cell sol. protein antigen (Ag) comprises the  
CC epitope to which human monoclonal antibody (mAb) 16-88 binds, found  
CC on a tumour associated Ag expressed by colon carcinoma cell lines HT-  
CC 29, S21463, SW948, SW403, LOVO and WiOr. Abs to the Ag may be used  
CC for the diagnosis and monitoring of treatment of cancer and in the  
CC prep. of vaccines to elicit an immune response. The Ag has the  
CC partial sequence given in R57740.  
SQ Sequence 14 AA;  
Query Match 26.3%; Score 5; DB 11; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 8 iqavr 12  
| | | | |  
QY 14 IQAVR 18  
RESULT 5  
ID P91703 standard; protein: 14 AA.  
AC P91703;  
DT 15-JUN-1990 (first entry)  
DE Human tumour cell epitope immunoreactive with human MAb 16-88.  
KW Cancer; colorectal cancer; hMAB 16-88.  
PN W08900050-A.  
PD 12-JAN-1989.  
PF 1-JUL-1988; 02245.  
PR 2-JUL-1987; US-069478.  
PA (ALKU) AKZO NV.  
PI Pomato N, Haspel MV, Hanna MG;  
DR WPI: 89-039545/05.  
PT Human tumor cell epitope -  
PT recognised by the human monoclonal antibody 16-88 used for  
PT producing antibodies for diagnosis and monitoring cancer.  
PS Claim 4; Page 26; 36pp; English.  
CC Partial sequence of human tumour cell epitope immunoreactive with human  
CC monoclonal antibody (hMAB) 16-88 capable of eliciting an immune response  
CC to tumour cells. Useful in diagnosis and monitoring tumour development.  
CC providing a vaccination against the tumour cells.  
CC See also P91704-5.  
SQ Sequence 14 AA;

Query Match 26.3%; Score 5; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 8 iqavr 12  
| | | | |  
QY 14 IQAVR 18  
RESULT 6  
ID W34092 standard; peptide: 20 AA.  
AC W34092;  
DT 05-MAY-1998 (first entry)  
DE Peptide derived from the N-terminal of the Presenilin-1 protein.  
KW Presenilin-1; antibody; cleavage; cleavage inhibition; treatment;  
KW prevention; Alzheimer's disease; cleavage assay; prognosis.  
OS Synthetic.  
OS Homo sapiens.  
PN W09741443-A2.  
PD 06-NOV-1997.  
PF 22-APR-1997; E02050.  
PR 26-APR-1996; GB-008657.  
PA (SMIK) SMITHKLINE BEECHAM PLC.  
PI Davis JB, Gray C, Karran EH, Ward RV;  
DR WPI: 97-549892/50.  
PT Identifying compounds that inhibit cleavage of presenilin-1 - for  
PT treatment, prevention and prognosis of Alzheimer's disease  
PS Claim 9; Page 15; 17pp; English.  
CC The present peptide is derived from the N-terminal of the presenilin-1  
CC protein, amino acid residues 42-60. Rabbit polyclonal antibodies were  
CC raised against this peptide. This antibody was used, together with an  
CC antibody raised against a peptide derived from the C-terminal, to  
CC demonstrate cleavage of presenilin-1 in cells. A novel method for  
CC screening compounds that modulate, especially inhibit, cleavage of  
CC presenilin-1 comprises measuring their effect on cleavage of presenilin-1  
CC to its 18 kD and 28 kD fragments. These compounds are potentially useful  
CC for treatment and prevention of Alzheimer's disease. Assays for cleavage  
CC can be used for prognosis of Alzheimer's disease.  
SQ Sequence 20 AA;  
Query Match 26.3%; Score 5; DB 28; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 10 lsngr 14  
| | | | |  
QY 9 LSNGR 13  
RESULT 7  
ID R57830 standard; protein: 25 AA.  
AC R57830;  
DT 22-MAR-1995 (first entry)  
DE Human tumour cell sol. protein antigen N-terminal.  
KW Tumour; cancer; soluble protein; antigen; antibody; Ab;  
KW monoclonal antibody; mAb; epitope; mAb 16-88;  
KW colon carcinoma cell line.  
OS Homo sapiens.  
PH Key Location/Qualifiers  
FT misc\_difference 15 /note= "undetermined amino acid"  
FT misc\_difference 17 /note= "undetermined amino acid"  
FT misc\_difference 18 /note= "undetermined amino acid"  
FT misc\_difference 20 /note= "undetermined amino acid"  
FT US5338832-A.  
PN 16-AUG-1994.  
PD 02-JUL-1987; 069478.  
PR 02-JUL-1987; US-069478.  
PR 28-FEB-1989; US-343475.

PR 13-AUG-1992; US-929842.

PA (ALKU ) AKZO NV.

PI Berman ML, Bos ES, Hanna MG, Haspel MV, Pomato N;

DR WPI; 94-263329/32.

PT Pure human tumour cell soluble protein antigen - comprises human

PT tumour cell epitope to which monoclonal antibody 16-88 binds

PS Disclosure; Col 7; 14pp; English.

CC The human tumour cell sol. protein antigen (Ag) comprises the

CC epitope to which human monoclonal antibody (mAb) 16-88 binds, found

CC on a tumour associated Ag expressed by colon carcinoma cell lines HT-

CC 29, S21463, SW948, SW403, LOVO and WiOr. Abs to the Ag may be used

CC for the diagnosis and monitoring of treatment of cancer and in the

CC prepn. of vaccines to elicit an immune response.

CC N-terminal protein sequence analysis was obtained from two bands

CC which were electroeluted after SDS-PAGE analysis. The N-terminal

CC ends of these bands are given in R57830-31.

SQ Sequence 25 AA;

Query Match 26.3%; Score 5; DB 11; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.81e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 iqavr 12

|||||

QY 14 IQAVR 18

RESULT 8

ID P91704 standard; protein; 25 AA.

AC P91704;

DT 15-JUN-1990 (first entry)

DE Human tumour cell epitope immunoreactive with human MAb 16-88.

KW Cancer; colorectal cancer; hMab 16-88.

PN W08900050-A.

PD 12-JAN-1989.

PF 1-JUL-1988; 02245.

PR 2-JUL-1987; US-069478.

PA (ALKU) AKZO NV.

PI Pomato N, Haspel MV, Hanna MG;

DR WPI; 89-039545/05.

PT Human tumour cell epitope -

PT recognised by the human monoclonal antibody 16-88 used for

PT producing antibodies for diagnosis and monitoring cancer.

PS Disclosure; 6pp; English.

CC N-terminal of subunit of human tumour cell epitope immunoreactive with

CC human monoclonal antibody (hMab) 16-88 capable of eliciting an immune

CC response to tumour cells. Useful in diagnosis and monitoring tumour

CC development, and providing a vaccination against the tumour cells.

CC See also P91703-5.

SQ Sequence 25 AA;

Query Match 26.3%; Score 5; DB 11; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.81e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 iqavr 12

|||||

QY 14 IQAVR 18

RESULT 9

ID R57831 standard; Protein; 26 AA.

AC R57831;

DT 22-MAR-1995 (first entry)

DE Human tumour cell sol. protein antigen N-terminal.

KW Tumour; cancer; soluble protein; antigen; antibody; Ab;

KW monoclonal antibody; mAb; epitope; mAb 16-88;

KW colon carcinoma cell line.

OS Homo sapiens.

FH key Location/Qualifiers

FT misc\_difference 17

/note= "undetermined amino acid"

FT misc\_difference 24

FT misc\_difference 25

/note= "undetermined amino acid"

FT misc\_difference 25

/note= "undetermined amino acid"

US5338832-A.

16-AUG-1994.

PD 02-JUL-1987; 069478.

PR 02-JUL-1987; US-069478.

PR 28-FEB-1989; US-343475.

PR 13-AUG-1992; US-929842.

PA (ALKU ) AKZO NV.

PI Berman ML, Bos ES, Hanna MG, Haspel MV, Pomato N;

DR WPI; 94-263329/32.

PT Pure human tumour cell soluble protein antigen - comprises human

PT tumour cell epitope to which monoclonal antibody 16-88 binds

PS Disclosure; Col 7; 14pp; English.

CC The human tumour cell sol. protein antigen (Ag) comprises the

CC epitope to which human monoclonal antibody (mAb) 16-88 binds, found

CC on a tumour associated Ag expressed by colon carcinoma cell lines HT-

CC 29, S21463, SW948, SW403, LOVO and WiOr. Abs to the Ag may be used

CC for the diagnosis and monitoring of treatment of cancer and in the

CC prepn. of vaccines to elicit an immune response.

CC N-terminal protein sequence analysis was obtained from two bands

CC which were electroeluted after SDS-PAGE analysis. The N-terminal

CC ends of these bands are given in R57830-31.

SQ Sequence 26 AA;

Query Match 26.3%; Score 5; DB 11; Length 26;

Best Local Similarity 100.0%; Pred. No. 1.81e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 iqavr 12

|||||

QY 14 IQAVR 18

RESULT 10

ID W50188 standard; peptide; 29 AA.

AC W50188;

DT 03-JUL-1998 (first entry)

DE Monoclonal antibody NTL Characterising peptide C.

KW Presenilin-1; NTL epitope; monoclonal antibody NTL;

KW Alzheimer's disease; diagnosis.

OS Homo sapiens.

PN W09804919-A1.

PD 05-FEB-1998.

PF 24-JUL-1997; U13374.

PR 25-JUL-1996; US-022594.

PA (MCLE-) MCLEAN HOSPITAL CORP.

PI Cataldo AM, Kao BH, Mathews PM, Nixon RA;

DR WPI; 98-130833/12.

PT Diagnosing Alzheimer's disease - from reduced levels of presenilin-1

PT and methods for screening compounds able to alter presenilin-1

PT levels, potentially useful for treatment of this disease

PS Example; Fig 1A; 3pp; English.

CC W50184 is the presenilin-1 (PS-1) NTL epitope, which was used in

CC the preparation of the anti-presenilin-1 monoclonal antibody NTL.

CC Alzheimer's disease (AD) is diagnosed by measuring the level of

CC PS-1, where a decreased level relative to controls indicates AD.

CC PS-1 is measured in lumbar or ventricular CSF, particularly by

CC standard immunoassays with NTL.

CC The coding region of PS-1 was amplified by PCR from a human liver

CC cDNA library using V2027/28, while the PCR primers V2029/30 were

CC used to generate a cDNA fragment encoding the PS-1 amino-terminal

CC fragment W50185. W50185 was expressed as a fusion protein with

CC glutathione-S-transferase, and used to raise NTL conventionally in

CC mice, which was characterised using the peptides W50186-88.

SQ Sequence 29 AA;

Query Match 26.3%; Score 5; DB 30; Length 29;

Best Local Similarity 100.0%; Pred. No. 1.81e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 lsng 15

```

CC as models for drug screening. The antibodies can also be used e.g. for
CC affinity purification and in immunoassays.
SQ Sequence 34 AA;

Query Match 26.3%; Score 5; DB 22; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.81e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 23 lsngr 27
QY 9 LSNGR 13

RESULT 13
ID R22388 standard; Protein: 40 AA.
AC R22388;
DT 17-AUG-1992 (first entry)
DE Antigen tc-35c.
KW Oocysts; antiserum; sporozoite; E. maxima.
OS Eimeria tenella.
PN W09204461-A.
PD 19-MAR-1992.
PF 05-SEP-1991; U06431.
PR 12-SEP-1990; US-581694.
PA (GENE-) GENEX CORP.
PA (FARH) HOECHST AG.
PI Jacobson JW, Strausberg RL, Wilson SD, Pope SH, Strausberg SL;
PI Raether W;
DR N-PSDB: Q23088.
PT Vaccine against avian coccidiosis - comprising recombinant
PT Eimeria antigen mc-4c, mc-5c or mc-30c gene, etc., or
PT microorganisms expressing them
PS Claim 24; Page 66 + Fig 24; 94pp; English.
CC To identify antigens of E. tenella, expression libraries were
CC prep'd. in the lambda vector, lambda gt11, using cDNA prep'd. from
CC polyA mRNA isolated from E. tenella oocysts. The cDNA expression
CC library was screened with rat anti-E. maxima sporozoite immune
CC serum. The cDNA inserts from the positive clones were cloned into
CC bacteriophage M13 and subjected to sequence analysis.
CC E. tenella antigen tc-35c (5 kD) was identified. Antigen tc-35c
CC consists of 120 bp of ORF encoding the carboxy terminal of an antigen.
SQ Sequence 40 AA;

Query Match 26.3%; Score 5; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.81e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 24 fylsn 28
QY 7 FYLSN 11

RESULT 14
ID W50185 standard; peptide: 47 AA.
AC W50185;
DT 03-JUL-1998 (first entry)
DE Presenilin-1 fragment.
KW Presenilin-1; NT1 epitope; monoclonal antibody NT1;
KW Alzheimer's disease; diagnosis.
OS Homo sapiens.
PN W09804919-A1.
PD 05-FEB-1998.
PF 24-JUL-1997; U13374.
PR 25-JUL-1996; US-022594.
PA (MCLE-) MCLEAN HOSPITAL CORP.
PI Cataldo AM, Kao BH, Mathews PM, Nixon RA;
PI WPI: 98-130833/12.
DR Diagnosing Alzheimer's disease - from reduced levels of presenilin-1
PT and methods for screening compounds able to alter presenilin-1
PT levels, potentially useful for treatment of this disease
PS Example; Fig 1A; 33pp; English.
CC W50184 is the presenilin-1 (PS-1) NT1 epitope, which was used in

```

```

QY 9 LSNGR 13

RESULT 11
ID P91705 standard; protein: 29 AA.
AC P91705;
DT 15-JUN-1990 (first entry)
DE Human tumour cell epitope immunoreactive with human MAB 16-88.
KW Cancer; colorectal cancer; hMAB 16-88.
PN W08900050-A.
PD 12-JAN-1989.
PF 1-JUL-1988; 02245.
PR 2-JUL-1987; US-069478.
PA (ALKU) Akzo NV.
PI Pomato N, Haspel MV, Hanna MG;
DR WPI: 89-039545/05.
PT Human tumour cell epitope -
PT recognised by the human monoclonal antibody 16-88 used for
PT producing antibodies for diagnosis and monitoring cancer.
PS Disclosure; 6pp; English.
CC N-terminal of subunit of human tumour cell epitope immunoreactive with
CC human monoclonal antibody (hMAB) 16-88 capable of eliciting an immune
CC response to tumour cells. Useful in diagnosis and monitoring tumour
CC development, and providing a vaccination against the tumour cells.
CC See also P91703-4.
SQ Sequence 29 AA;

Query Match 26.3%; Score 5; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.81e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 iqavr 12
QY 14 IQAVR 18

RESULT 12
ID W05769 standard; peptide: 34 AA.
AC W05769;
DT 28-JUL-1997 (first entry)
DE Presenilin-1 residues 28-61.
KW Presenilin-1; human; hPS1-1; hPS1-2; PS-2; integral membrane protein; AD;
KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
KW depression; antibody; gene expression modulator; therapy; muten.
OS Homo sapiens.
PN W09634099-A2.
PD 31-OCT-1996.
PF 29-APR-1996; CA0263.
PR 28-APR-1995; US-431048.
PR 28-JUN-1995; US-496841.
PR 31-JUL-1995; US-509359.
PA (HSCR) HSC RES & DEV LP.
PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
PI Fraser PE, Rommens JM, St George-Hyslop PH;
DR WPI: 96-497631/49.
PT New presenilin genes - useful for diagnosis, therapy and drug
PT screening of familial Alzheimer's disease, cerebral disorders, etc.
PS Claim 71; Page -; 178pp; English.
CC W05768-W05788 represent antigenic fragments of the human presenilin-1-1
CC protein (see W05733 for wild type sequence). W05734 represents a
CC different wild type form of presenilin-1 that results from alternate
CC splicing of the genomic DNA sequence. The presenilins are a family of
CC highly conserved integral membrane proteins with a common structural
CC motif, common alternate splicing patterns, and common mutational hot spot
CC regions. Mutations in PS genes are implicated in familial Alzheimer's
CC disease (AD) and possibly other diseases such as cerebral haemorrhage,
CC schizophrenia, depression etc., so detection of mutations in the DNA
CC encoding the wild type sequences can be used for diagnosis of these
CC diseases. The wild type proteins, or vectors that express them or
CC containing antisense sequences, antibodies selective for these mutant
CC forms of the proteins and modulators of PS gene expression are
CC potentially useful for treatment of AD etc. Transgenic animals are useful

```

CC the preparation of the anti-presenilin-1 monoclonal antibody NT1.  
 CC Alzheimer's disease (AD) is diagnosed by measuring the level of  
 CC PS-1, where a decreased level relative to controls indicates AD.  
 CC PS-1 is measured in lumbar or ventricular CSF, particularly by  
 CC standard immunoassays with NT1.  
 CC The coding region of PS-1 was amplified by PCR from a human liver  
 CC cDNA library using V2027/28, while the PCR primers V2029/30 were  
 CC used to generate a cDNA fragment encoding the PS-1 amino-terminal  
 CC fragment W50185. W50185 was expressed as a fusion protein with  
 CC glutathione-S-transferase, and used to raise NT1 conventionally in  
 CC mice, which was characterised using the peptides W50186-88.  
 SQ Sequence 47 AA;

Query Match 26.3%; Score 5; DB 30; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 30 lsngr 34  
 |||||  
 QY 9 LSNGR 13

RESULT 15  
 ID W57076 standard; peptide: 60 AA.  
 AC W57076;  
 DT 30-JUL-1998 (first entry)  
 DE Measles virus haemagglutinin neutralising epitope region MEAHAA.  
 KW Measles virus; haemagglutinin neutralising epitope region; MV isolate;  
 KW morbillivirus; vaccine; immune response; Edmonston strain; HNE;  
 KW antigenic.  
 OS Synthetic.  
 OS Measles virus.  
 PN W09805682-A2.  
 PD 12-FEB-1998.  
 PF 26-JUL-1997; E04070.  
 PR 31-JUL-1996; EP-112341.  
 PA (MULL/) MULLER C P.  
 PI Muller CP;  
 DR WPI: 98-145549/13.  
 PT Morbillivirus antigenic determinant peptides - useful for, e.g.  
 PT preparation of protective vaccines against measles  
 PS Claim 8; Fig 7; 67pp; English.  
 CC The present sequence represents a peptide of the haemagglutinin  
 CC neutralising epitope region from a measles virus (MV) isolate. The  
 CC peptide has an amino acid sequence corresponding to, or mimicking  
 CC at least 1 antigenic determinant of the morbillivirus haemagglutinin  
 CC protein. The peptide can be used to provoke an immune response to  
 CC measles virus in mammals, and is useful in the preparation of a  
 CC vaccine for the prevention of measles. The peptide is present in a  
 CC vaccine at about 0.1 mu g to 100 mu g.  
 SQ Sequence 60 AA;

Query Match 26.3%; Score 5; DB 31; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 1.81e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 28 griqa 32  
 |||||  
 QY 12 GRIQA 16

Search completed: Sat Aug 28 14:41:32 1999  
 Job time : 76 secs.

\*\*\*\*\*  
W O R L D  
\*\*\*\*\*  
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Aug 28 14:39:13 1999; Maspar time 4.80 Seconds  
Tabular output not generated.  
158.532 Million cell updates/sec

Title: >US-09-049-696-42  
Description: (1-19) from US09049696.ppep  
Perfect Score: 19  
Sequence: 1 YNNDEKPYLSNGRIQAVRC 19

Scoring table: TABLE unitprotable  
Gap 60

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: pir60  
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 2.498; Variance 0.411; scale 6.080

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	10	52.6	913	3	JG0168	8.88e-08
2	7	36.8	2573	2	D71614	hypothetical protein
3	6	31.6	97	2	G69121	hypothetical protein
4	6	31.6	151	2	G37812	hypothetical protein
5	6	31.6	181	2	S73534	hypothetical protein
6	6	31.6	213	2	S58329	HS49 protein - yeast
7	6	31.6	314	2	S75995	hypothetical protein
8	6	31.6	316	2	S77783	hypothetical protein
9	6	31.6	332	2	S71224	xyloglucan endo-1,4-b
10	6	31.6	356	2	S32147	naringenin 3-dioxygen
11	6	31.6	374	2	S77750	naringenin 3-dioxygen
12	6	31.6	410	1	JN0810	benzene 1,2-dioxyg
13	6	31.6	458	2	G69123	conserved hypoch
14	6	31.6	545	2	F64579	adenine specific DNA
15	6	31.6	545	2	F71932	type II DNA modif
16	6	31.6	547	2	A64991	hypothetical ABC tran
17	6	31.6	574	2	B35149	ipah protein - Shigel
18	6	31.6	621	2	T01935	naringenin 3-dioxygen
19	6	31.6	1009	2	S28857	glutamate receptor de
20	6	31.6	1009	2	JH0266	glutamate receptor de
21	6	31.6	1451	1	QJ7119	E2 glycoprotein precu
22	5	26.3	58	2	C27606	Ig heavy chain V-a re
23	5	26.3	62	2	I36968	gene MHC DQ-alpha 1 p

gene MHC DQ-alpha 1 p  
endonuclease precursor  
dopamine D3 receptor  
pulmonary surfactant  
C14L protein - vacci  
hypothetical 8.8K pro  
hypothetical protein  
hypothetical 10.5K pr  
ribosomal protein S14  
hypothetical protein  
single-strand DNA-bin  
gene 60.1 protein - p  
hypothetical protein  
allene oxide synthase  
hypothetical protein  
hypothetical protein  
hypothetical protein  
transhyretin precurs  
hypothetical protein  
prrD protein homolog  
hypothetical protein  
mannose-binding lecti  
ribosomal protein L21  
mobilization protein  
hypothetical protein  
mannose-binding lecti  
capsid protein - oliv  
ferric uptake regulat  
trbB protein precursor  
probable lpgE protein  
hypothetical protein  
fix23-4 protein - Rhl  
hypothetical protein  
21.5K protein - vacci  
hypothetical protein  
hypothetical protein  
Achaete-scute complex  
gene achaete protein  
endothelin 1 precursor  
protein YNR023w homol  
N-methyl-D-aspartate  
hypothetical protein  
regulatory protein cy  
vifN-28 protein - fava  
hypothetical protein  
conserved hypochetica  
hypothetical protein  
signal-transducing pr  
transcription initiat  
hypothetical protein  
serine O-acetyltransf  
hypothetical protein  
serine O-acetyltransf  
serine O-acetyltransf  
C08C3.2 protein - Cae  
ubiquinol oxidase (EC  
hypothetical protein  
hypothetical protein  
nodulin-30 precursor  
hypothetical 31.5 kD  
nodulin-31 precursor  
ykfa protein - Escher  
yfpj protein - Escher  
xyloglucan endo-1,4-b  
nodulin-32b precursor  
nodulin-32a precursor  
translation elongatio  
transcription initiat  
matrix protein - suba  
matrix protein - meas  
matrix protein - meas  
matrix protein - meas

24 5 26.3 63 2 I36901  
25 5 26.3 72 2 S53874  
26 5 26.3 75 2 A55419  
27 5 26.3 79 1 LNPG1  
28 5 26.3 82 2 I42502  
29 5 26.3 83 2 JN0400  
30 5 26.3 86 2 G64163  
31 5 26.3 93 2 JE0123  
32 5 26.3 100 2 A27523  
33 5 26.3 108 2 D71008  
34 5 26.3 113 2 B70066  
35 5 26.3 126 2 JS0540  
36 5 26.3 129 2 S77218  
37 5 26.3 132 2 A69314  
38 5 26.3 140 2 E71235  
39 5 26.3 142 2 S05662  
40 5 26.3 142 2 S75218  
41 5 26.3 149 2 S67473  
42 5 26.3 161 2 S40711  
43 5 26.3 163 2 F40555  
44 5 26.3 164 2 A70538  
45 5 26.3 165 2 S38257  
46 5 26.3 166 2 A48465  
47 5 26.3 172 2 JQ0391  
48 5 26.3 174 2 B71453  
49 5 26.3 174 2 S38258  
50 5 26.3 175 2 S42095  
51 5 26.3 176 2 F70180  
52 5 26.3 179 1 BVECTB  
53 5 26.3 182 2 A70804  
54 5 26.3 183 2 S28803  
55 5 26.3 188 2 S18956  
56 5 26.3 188 2 S12393  
57 5 26.3 190 1 WVVZ21  
58 5 26.3 192 2 S75403  
59 5 26.3 198 2 S01445  
60 5 26.3 201 2 A43731  
61 5 26.3 201 2 S23324  
62 5 26.3 202 1 S63538  
63 5 26.3 209 2 S61906  
64 5 26.3 221 2 I51308  
65 5 26.3 233 2 G70776  
66 5 26.3 236 2 A30225  
67 5 26.3 238 2 S42657  
68 5 26.3 248 2 T02554  
69 5 26.3 250 2 E69851  
70 5 26.3 251 2 B71427  
71 5 26.3 254 2 D64540  
72 5 26.3 255 2 C37165  
73 5 26.3 257 2 A69996  
74 5 26.3 267 2 G64080  
75 5 26.3 267 2 G71069  
76 5 26.3 273 1 XVECSA  
77 5 26.3 273 2 S29568  
78 5 26.3 274 2 S44616  
79 5 26.3 277 2 JC5900  
80 5 26.3 278 2 S23355  
81 5 26.3 278 2 C70138  
82 5 26.3 279 2 S42654  
83 5 26.3 282 2 H51116  
84 5 26.3 287 2 S42655  
85 5 26.3 288 2 E64750  
86 5 26.3 289 2 B65042  
87 5 26.3 295 2 S48102  
88 5 26.3 297 2 S42656  
89 5 26.3 298 2 S42653  
90 5 26.3 298 2 H64247  
91 5 26.3 300 2 S60164  
92 5 26.3 300 1 B60285  
93 5 26.3 302 2 P00383  
94 5 26.3 302 2 P00379  
95 5 26.3 302 2 P00375  
96 5 26.3 302 2 P00387

97	5	26.3	305	2	I41065	glycine cleavage syst	1.05e-02	170	5	26.3	440	2	C70198	conserved hypothetical	1.05e+02
98	5	26.3	305	2	S40927	hypothetical protein	1.05e-02	171	5	26.3	441	2	E70075	monooxygenase homolog	1.05e-02
99	5	26.3	305	2	S43966	preR protein - Pseudo	1.05e-02	172	5	26.3	442	2	A71969	probable histidine ki	1.05e-02
100	5	26.3	310	2	S43865	cytochrome P-450 2A1	1.05e-02	173	5	26.3	443	2	C70444	biotin carboxylase	1.05e-02
101	5	26.3	312	2	B55461	NAD(P)+--arginine ADP	1.05e-02	174	5	26.3	444	2	A25539	O-acetyl homoserine-O	1.05e-02
102	5	26.3	318	2	S27977	collagen dpy-7 - Caen	1.05e-02	175	5	26.3	445	2	A45139	oligosaccharyltransfe	1.05e-02
103	5	26.3	321	2	D59101	coenzyme F420-depende	1.05e-02	176	5	26.3	446	2	E70854	probable membrane pro	1.05e-02
104	5	26.3	321	2	S66259	N5,N10-methylenetetra	1.05e-02	177	5	26.3	446	2	DYRTD3	dopamine receptor D3	1.05e-02
105	5	26.3	323	2	S72476	nodulin 28/32 precurs	1.05e-02	178	5	26.3	446	2	148322	dopamine receptor D3	1.05e-02
106	5	26.3	328	1	F64187	p-aminobenzoate synth	1.05e-02	179	5	26.3	449	2	S28167	cytochrome P450 2E1 -	1.05e-02
107	5	26.3	329	1	NBXYN1	nuclease NUC1 (EC 3.1	1.05e-02	180	5	26.3	455	2	S50725	hypothetical protein	1.05e-02
108	5	26.3	331	2	B45451	beta-ketoacyl-acyl ca	1.05e-02	181	5	26.3	460	2	A43563	ornithine decarboxyla	1.05e-02
109	5	26.3	331	2	C171962	beta-ketoacyl-acyl syn	1.05e-02	182	5	26.3	462	2	A40362	anthranilate synthase	1.05e-02
110	5	26.3	335	2	S11227	matrix protein - meas	1.05e-02	183	5	26.3	463	2	S63683	presenilin 1-463 - hu	1.05e-02
111	5	26.3	335	1	MFN2MH	gene M protein - rind	1.05e-02	184	5	26.3	463	2	JC5081	presenilin 1 protein	1.05e-02
112	5	26.3	335	1	MFN2RP	matrix protein - rind	1.05e-02	185	5	26.3	464	2	C69088	anthranilate synthase	1.05e-02
113	5	26.3	335	1	MFN2MV	matrix protein - meas	1.05e-02	186	5	26.3	467	2	S58396	presenilin 1, splice	1.05e-02
114	5	26.3	335	2	S35933	matrix protein - meas	1.05e-02	187	5	26.3	467	2	JC5080	presenilin 1 protein	1.05e-02
115	5	26.3	335	1	D48556	matrix protein - meas	1.05e-02	188	5	26.3	475	2	C71404	hypothetical protein	1.05e-02
116	5	26.3	335	1	JQ1566	matrix protein - phoc	1.05e-02	189	5	26.3	483	2	A34720	keratin 8, type II cy	1.05e-02
117	5	26.3	335	1	MFN2MH	matrix protein - meas	1.05e-02	190	5	26.3	485	2	B69795	amidase homolog verM	1.05e-02
118	5	26.3	335	2	B49601	matrix protein M - me	1.05e-02	191	5	26.3	487	2	F70765	hypothetical protein	1.05e-02
119	5	26.3	335	1	MFN2CV	matrix protein - can	1.05e-02	192	5	26.3	487	2	JT0407	keratin 8, type II cy	1.05e-02
120	5	26.3	336	2	S35934	matrix protein - meas	1.05e-02	193	5	26.3	489	2	S05474	keratin 8, type II, c	1.05e-02
121	5	26.3	336	2	S35935	matrix protein - meas	1.05e-02	194	5	26.3	490	2	JC5068	cytochrome P450 2E1 -	1.05e-02
122	5	26.3	339	2	A70321	histidine kinase sens	1.05e-02	195	5	26.3	493	2	A31949	cytochrome P450 2E1 -	1.05e-02
123	5	26.3	343	1	MFN2BK	matrix protein - meas	1.05e-02	196	5	26.3	497	2	S37581	probable acyltransfer	1.05e-02
124	5	26.3	345	2	B42604	ORF2 complementary to	1.05e-02	197	5	26.3	497	2	S37580	probable acyltransfer	1.05e-02
125	5	26.3	347	2	S40082	hypothetical protein	1.05e-02	198	5	26.3	498	2	S20918	probable serine/threo	1.05e-02
126	5	26.3	347	2	A38453	anaerobic sulfite red	1.05e-02	199	5	26.3	505	2	S64837	hypothetical protein	1.05e-02
127	5	26.3	349	2	D70398	hypothetical protein	1.05e-02	200	5	26.3	505	2	JC5072	cytochrome P450 ALK4,	1.05e-02
128	5	26.3	349	2	D64134	peptide transport sys	1.05e-02	201	5	26.3	507	2	S05542	hypothetical protein,	1.05e-02
129	5	26.3	352	2	S70972	bfpE protein - Escher	1.05e-02	202	5	26.3	509	2	S74935	hypothetical protein,	1.05e-02
130	5	26.3	354	2	E69035	sensory transduction	1.05e-02	203	5	26.3	510	2	JC2311	NADH oxidase (hydroge	1.05e-02
131	5	26.3	355	2	S47707	hypothetical 38.8K pr	1.05e-02	204	5	26.3	511	2	B34160	cytochrome P450 4A7 -	1.05e-02
132	5	26.3	357	2	S38338	naringenin 3-dioxygen	1.05e-02	205	5	26.3	514	2	D56849	dopamine receptor-lik	1.05e-02
133	5	26.3	357	2	S71772	naringenin 3-dioxygen	1.05e-02	206	5	26.3	515	2	PC4419	actin-binding 260K pr	1.05e-02
134	5	26.3	357	2	S61415	naringenin 3-dioxygen	1.05e-02	207	5	26.3	516	2	FWCNBB	beta-globulin B precu	1.05e-02
135	5	26.3	363	1	BYBYK2	MAK32 protein - yeast	1.05e-02	208	5	26.3	521	2	JC5073	cytochrome P450 ALK5-	1.05e-02
136	5	26.3	363	2	S74518	hypothetical protein	1.05e-02	209	5	26.3	522	2	C64139	peptide-chain-release	1.05e-02
137	5	26.3	364	2	S31458	naringenin 3-dioxygen	1.05e-02	210	5	26.3	534	2	S70971	bfpD protein - Escher	1.05e-02
138	5	26.3	365	2	S31921	naringenin 3-dioxygen	1.05e-02	211	5	26.3	539	2	T02508	myosin heavy chain-li	1.05e-02
139	5	26.3	366	2	S70674	bplC protein - Bordet	1.05e-02	212	5	26.3	549	2	F64640	conserved hypothetical	1.05e-02
140	5	26.3	367	2	N06634	caricain (EC 3.4.22.3	1.05e-02	213	5	26.3	550	2	A48026	sterol O-acyltransfer	1.05e-02
141	5	26.3	368	2	G70370	conserved hypothetical	1.05e-02	214	5	26.3	555	2	S12639	ATP-binding protein p	1.05e-02
142	5	26.3	369	2	S16780	naringenin 3-dioxygen	1.05e-02	215	5	26.3	562	2	S56145	BS69 protein - human	1.05e-02
143	5	26.3	369	2	A42110	flavanone 3 beta-hydr	1.05e-02	216	5	26.3	565	2	S27778	merozoite surface ant	1.05e-02
144	5	26.3	371	2	S30978	integrase - Mycobacte	1.05e-02	217	5	26.3	566	2	S19063	hypothetical protein	1.05e-02
145	5	26.3	372	2	T03385	naringenin 3-dioxygen	1.05e-02	218	5	26.3	575	2	S11485	RNA helicase - fruit	1.05e-02
146	5	26.3	374	2	S63684	presenilin 1, splice	1.05e-02	219	5	26.3	580	3	JH0224	site-specific methylt	1.05e-02
147	5	26.3	379	2	T01749	gibberellin 20-oxidas	1.05e-02	220	5	26.3	581	2	S04857	penicillin-binding pr	1.05e-02
148	5	26.3	379	2	T01748	gibberellin 20-oxidas	1.05e-02	221	5	26.3	581	2	S00916	penicillin-binding pr	1.05e-02
149	5	26.3	380	2	JN0594	pectate lyase (EC 4.2	1.05e-02	222	5	26.3	581	2	S49096	penicillin-binding pr	1.05e-02
150	5	26.3	381	2	F48833	sulfate starvation-in	1.05e-02	223	5	26.3	581	2	S49093	penicillin-binding pr	1.05e-02
151	5	26.3	383	1	A25978	alcohol dehydrogenase	1.05e-02	224	5	26.3	581	2	S49098	penicillin-binding pr	1.05e-02
152	5	26.3	386	2	E71693	hypothetical protein	1.05e-02	225	5	26.3	582	2	S51945	ferredoxin--nitrite r	1.05e-02
153	5	26.3	387	2	S52274	transcription termina	1.05e-02	226	5	26.3	582	2	S49097	penicillin-binding pr	1.05e-02
154	5	26.3	387	1	ERADN1	41K fiber protein - h	1.05e-02	227	5	26.3	582	2	A36190	penicillin-binding pr	1.05e-02
155	5	26.3	389	2	S31748	heat shock protein - h	1.05e-02	228	5	26.3	582	2	S49095	penicillin-binding pr	1.05e-02
156	5	26.3	395	2	S12043	gene A2 protein - mai	1.05e-02	229	5	26.3	582	2	S49091	penicillin-binding pr	1.05e-02
157	5	26.3	400	2	G01977	d3 dopamine receptor	1.05e-02	230	5	26.3	582	2	S49092	penicillin-binding pr	1.05e-02
158	5	26.3	400	2	G00013	D3 dopamine receptor	1.05e-02	231	5	26.3	582	2	S49090	penicillin-binding pr	1.05e-02
159	5	26.3	402	2	S55980	probable membrane pro	1.05e-02	232	5	26.3	586	2	JN0505	protein kinase (EC 2.	1.05e-02
160	5	26.3	403	2	S35334	steroid receptor prot	1.05e-02	233	5	26.3	588	1	BYEGUC	excinuclease ABC, cha	1.05e-02
161	5	26.3	403	2	A25334	sex steroid-binding p	1.05e-02	234	5	26.3	591	2	B71920	hypothetical protein	1.05e-02
162	5	26.3	406	2	C48002	N-acetylglucosamine r	1.05e-02	235	5	26.3	591	2	I38552	aspartic proteinase M	1.05e-02
163	5	26.3	410	1	D36516	toluene dioxygenase (	1.05e-02	236	5	26.3	596	2	S57971	probable DNA primase	1.05e-02
164	5	26.3	417	2	S64245	hypothetical protein	1.05e-02	237	5	26.3	600	2	S76764	hypothetical protein	1.05e-02
165	5	26.3	423	1	XPXS2M	dihydrodipicamide S-(2	1.05e-02	238	5	26.3	605	2	D71318	probable DNA primase	1.05e-02
166	5	26.3	425	2	F71871	topoisomerase I - Hel	1.05e-02	239	5	26.3	606	2	A33496	sensory kinase (EC 2.	1.05e-02
167	5	26.3	428	2	A55044	beta-4C-adrenergic re	1.05e-02	240	5	26.3	609	2	F64045	excinuclease ABC chal	1.05e-02
168	5	26.3	428	2	D46615	hypothetical protein	1.05e-02	241	5	26.3	609	2	A30311	protein kinase C (EC	1.05e-02
169	5	26.3	437	2	C35147	integrase homolog - S	1.05e-02	242	5	26.3	615	2	H64769	preprotein translocas	1.05e-02



243	5	26.3	621	2	138880	eleven-nineteen lysin	1.05e+02	316	5	26.3	1145	2	S13643	PRP22 protein - yeast	1.05e+02
244	5	26.3	622	2	S12981	68K protein - gray wo	1.05e+02	317	5	26.3	1155	2	H71456	probable pyrolysin -	1.05e+02
245	5	26.3	627	2	JH0595	gamma-aminobutyric ac	1.05e+02	318	5	26.3	1193	2	G71605	hypothetical protein	1.05e+02
246	5	26.3	627	2	B44099	gamma-aminobutyric ac	1.05e+02	319	5	26.3	1251	2	G49645	probable membrane pro	1.05e+02
247	5	26.3	635	2	S27714	csfB protein - Clostr	1.05e+02	320	5	26.3	1262	2	B48758	protein-tyrosine-phos	1.05e+02
248	5	26.3	635	2	S73017	polyketide synthase p	1.05e+02	321	5	26.3	1432	2	S58819	antiviral protein SKI	1.05e+02
249	5	26.3	643	2	S49303	dnak-type molecular c	1.05e+02	322	5	26.3	1446	2	S73013	polyketide synthase p	1.05e+02
250	5	26.3	649	2	T01882	hypothetical protein	1.05e+02	323	5	26.3	1496	2	A48758	protein-tyrosine-phos	1.05e+02
251	5	26.3	649	2	S16682	methionine--trna liga	1.05e+02	324	5	26.3	1499	2	I50212	protein-tyrosine-phos	1.05e+02
252	5	26.3	679	2	G711615	phospholipase A2-like	1.05e+02	325	5	26.3	1501	2	I58148	protein-tyrosine-phos	1.05e+02
253	5	26.3	679	2	S28366	recombination repair	1.05e+02	326	5	26.3	1501	2	S46218	protein-tyrosine-phos	1.05e+02
254	5	26.3	692	2	I61777	Munc13-3 - rat (fragm	1.05e+02	327	5	26.3	1613	2	A43081	vitellogenin 2 precu	1.05e+02
255	5	26.3	696	2	A70373	(p)ppgpp 3-pyrophosph	1.05e+02	328	5	26.3	1694	2	S50065	sialoadhesin - mouse	1.05e+02
256	5	26.3	700	2	S38426	chaperonin 60 precurs	1.05e+02	329	5	26.3	1739	2	A48298	sodium channel homolo	1.05e+02
257	5	26.3	703	2	S67277	Probable membrane pro	1.05e+02	330	5	26.3	1769	2	S53378	probable membrane pro	1.05e+02
258	5	26.3	714	2	C64239	ATP-dependent protein	1.05e+02	331	5	26.3	1863	2	S46217	protein-tyrosine-phos	1.05e+02
259	5	26.3	736	1	S31809	translation elongatio	1.05e+02	332	5	26.3	1897	1	TDHULK	leukocyte antigen-rel	1.05e+02
260	5	26.3	742	2	T00371	hypothetical protein	1.05e+02	333	5	26.3	1898	2	S46216	leukocyte antigen-rel	1.05e+02
261	5	26.3	745	2	S48019	kinesin-related prote	1.05e+02	334	5	26.3	1907	2	S50893	protein-tyrosine-phos	1.05e+02
262	5	26.3	759	2	F64652	trans-Golgi membrane	1.05e+02	335	5	26.3	2010	2	B71616	phosphatase (acid pho	1.05e+02
263	5	26.3	759	2	D71853	hypothetical protein	1.05e+02	336	5	26.3	2039	1	TDFFLK	protein-tyrosine-phos	1.05e+02
264	5	26.3	768	2	JC6564	cellulose oxidase (E	1.05e+02	337	5	26.3	2319	2	A47004	coagulation factor VI	1.05e+02
265	5	26.3	775	2	B64319	carbox-monoxide dehyd	1.05e+02	338	5	26.3	2335	2	A61208	chondroitin sulfate p	1.05e+02
266	5	26.3	778	2	T00412	outer membrane protei	1.05e+02	339	5	26.3	3343	2	S44887	zk112.7 protein - Cae	1.05e+02
267	5	26.3	781	2	T00456	probable protein kina	1.05e+02	340	5	26.3	4447	2	A69679	polyketide synthase p	1.05e+02
268	5	26.3	796	2	JC1285	protein-tyrosine-phos	1.05e+02	341	4	21.1	15	2	S03955	acidic fibroblast gro	2.35e+03
269	5	26.3	802	2	A36065	protein-tyrosine-phos	1.05e+02	342	4	21.1	33	2	S32764	T-cell receptor gamma	2.35e+03
270	5	26.3	829	2	A47373	protein-tyrosine-phos	1.05e+02	343	4	21.1	34	2	S35923	T-cell receptor gamma	2.35e+03
271	5	26.3	830	2	S25198	vacuolar membrane pro	1.05e+02	344	4	21.1	57	2	E42075	finger protein (clone	2.35e+03
272	5	26.3	853	2	D70304	hypothetical protein	1.05e+02	345	4	21.1	58	2	S28251	NADH dehydrogenase (u	2.35e+03
273	5	26.3	857	2	S01864	lipoygenase (EC 1.13	1.05e+02	346	4	21.1	67	2	I69230	CE5 - dog	2.35e+03
274	5	26.3	861	2	S01442	lipoygenase (EC 1.13	1.05e+02	347	4	21.1	74	2	I48263	phosphatidylinositol-	2.35e+03
275	5	26.3	863	2	B69301	signal-transducing hi	1.05e+02	348	4	21.1	78	2	S61473	p83/100 protein - Bor	2.35e+03
276	5	26.3	875	1	URX1A2	peptidylglycine monoo	1.05e+02	349	4	21.1	78	2	S61474	p83/100 protein - Bor	2.35e+03
277	5	26.3	875	2	S66672	phosphatidylinositol-	1.05e+02	350	4	21.1	78	2	S61475	p83/100 protein - Bor	2.35e+03
278	5	26.3	879	2	S55864	hypothetical protein	1.05e+02	351	4	21.1	83	2	S24712	Ig alpha chain - huma	2.35e+03
279	5	26.3	882	2	A39030	androgen-binding prot	1.05e+02	352	4	21.1	85	2	G69919	hypothetical protein	2.35e+03
280	5	26.3	882	2	FC1231	hypothetical protein	1.05e+02	353	4	21.1	85	2	A42075	finger protein (clone	2.35e+03
281	5	26.3	885	2	JN0339	N-methyl-D-aspartate	1.05e+02	354	4	21.1	87	2	T03360	gene e7 protein - lac	2.35e+03
282	5	26.3	885	2	A47551	N-methyl-D-aspartate	1.05e+02	355	4	21.1	90	2	S54726	STAR5 protein homolog	2.35e+03
283	5	26.3	891	2	S75940	methyl-accepting chem	1.05e+02	356	4	21.1	91	2	H69543	NADH dehydrogenase ch	2.35e+03
284	5	26.3	892	2	A56213	DNA excision-repair p	1.05e+02	357	4	21.1	91	2	S22894	T-cell receptor alpha	2.35e+03
285	5	26.3	901	2	JN0337	N-methyl-D-aspartate	1.05e+02	358	4	21.1	96	2	A40081	R87 antigen - rat	2.35e+03
286	5	26.3	906	2	JN0341	N-methyl-D-aspartate	1.05e+02	359	4	21.1	104	2	S53260	hypothetical protein	2.35e+03
287	5	26.3	906	2	A46296	N-methyl-D-aspartate	1.05e+02	360	4	21.1	105	2	D71152	hypothetical protein	2.35e+03
288	5	26.3	911	2	S28098	ste6 protein - fissio	1.05e+02	361	4	21.1	108	2	S22892	T-cell receptor alpha	2.35e+03
289	5	26.3	922	2	JN0338	N-methyl-D-aspartate	1.05e+02	362	4	21.1	111	2	S22771	B1496_FL_23 protein -	2.35e+03
290	5	26.3	922	2	JN0340	N-methyl-D-aspartate	1.05e+02	363	4	21.1	114	2	S69320	probable membrane pro	2.35e+03
291	5	26.3	923	2	S09583	peptidylglycine monoo	1.05e+02	364	4	21.1	116	2	JE0355	pheromone binding pro	2.35e+03
292	5	26.3	935	2	S17855	peptidylglycine monoo	1.05e+02	365	4	21.1	116	2	JE0354	pheromone binding pro	2.35e+03
293	5	26.3	938	2	S21104	N-methyl-D-aspartate	1.05e+02	366	4	21.1	117	2	F71065	hypothetical protein	2.35e+03
294	5	26.3	938	2	A46612	N-methyl-D-aspartate	1.05e+02	367	4	21.1	118	2	E71158	hypothetical protein	2.35e+03
295	5	26.3	938	2	S19710	N-methyl-D-aspartate	1.05e+02	368	4	21.1	120	2	T02708	hypothetical protein	2.35e+03
296	5	26.3	943	2	S59317	DIP2 protein - yeast	1.05e+02	369	4	21.1	130	2	S08496	hypothetical protein	2.35e+03
297	5	26.3	950	2	S69704	hypothetical protein	1.05e+02	370	4	21.1	132	2	H71273	probable ribosomal pr	2.35e+03
298	5	26.3	951	1	E64066	ribonuclease E (EC 3.	1.05e+02	371	4	21.1	134	2	C48234	hypothetical protein	2.35e+03
299	5	26.3	958	1	PIBVCC	la protein - cowpea c	1.05e+02	372	4	21.1	134	2	S08640	hypothetical protein	2.35e+03
300	5	26.3	959	2	JN0336	N-methyl-D-aspartate	1.05e+02	373	4	21.1	136	2	A53260	hypothetical protein	2.35e+03
301	5	26.3	965	2	I51244	N-methyl-D-aspartate	1.05e+02	374	4	21.1	134	4	C44264	ALL-1/AF-4 clone 25 m	2.35e+03
302	5	26.3	971	2	A70179	exodeoxyribonuclease	1.05e+02	375	4	21.1	135	3	JE0310	serine protease inhib	2.35e+03
303	5	26.3	972	1	URBOAP	peptidylglycine monoo	1.05e+02	376	4	21.1	136	2	C71018	probable NADH-plastoq	2.35e+03
304	5	26.3	974	1	URHUAP	peptidylglycine monoo	1.05e+02	377	4	21.1	138	2	F69499	hypothetical protein	2.35e+03
305	5	26.3	976	1	URRTAP	kinase-like protein k	1.05e+02	378	4	21.1	138	2	S51556	K222 protein - human	2.35e+03
306	5	26.3	987	2	A64474	hypothetical protein	1.05e+02	379	4	21.1	142	2	JC6305	neutrin precursor -	2.35e+03
307	5	26.3	1001	2	H64593	type III restriction	1.05e+02	380	4	21.1	143	2	B71159	hypothetical protein	2.35e+03
308	5	26.3	1034	1	A53653	enteropeptidase (EC 3	1.05e+02	381	4	21.1	144	2	F70417	hypothetical protein	2.35e+03
309	5	26.3	1050	2	S45636	natriuretic-peptide r	1.05e+02	382	4	21.1	151	2	G70395	hypothetical protein	2.35e+03
310	5	26.3	1051	2	A39712	kinase-like protein k	1.05e+02	383	4	21.1	152	2	B71035	hypothetical protein	2.35e+03
311	5	26.3	1051	2	S59791	probable membrane pro	1.05e+02	384	4	21.1	154	2	A36144	cobE protein - Pseudo	2.35e+03
312	5	26.3	1102	2	S28104	probable DNA-directed	1.05e+02	385	4	21.1	157	2	S58073	probable olfactory re	2.35e+03
313	5	26.3	1113	2	S48495	probable membrane pro	1.05e+02	386	4	21.1	158	2	T03330	endodeoxyribonuclease	2.35e+03
314	5	26.3	1136	2	C71110	hypothetical protein	1.05e+02	387	4	21.1	161	2	E71099	hypothetical protein	2.35e+03
315	5	26.3	1139	2	D64503	hypothetical protein	1.05e+02	388	4	21.1	168	2	S62881	calcium binding prote	2.35e+03

389	4	21.1	170	2	B71864	hypothetical protein	2.35e+03	462	4	21.1	281	2	S15191	gvpL protein - Haloba	2.35e+03
390	4	21.1	170	2	H71454	hypothetical protein	2.35e+03	463	4	21.1	281	2	JQ1129	GvpL protein - Haloba	2.35e+03
391	4	21.1	171	2	C71244	hypothetical protein	2.35e+03	464	4	21.1	284	2	S93680	hypothetical protein	2.35e+03
392	4	21.1	174	2	S39790	male-enhanced antigen	2.35e+03	465	4	21.1	285	2	S93288	TIS11 protein - yeast	2.35e+03
393	4	21.1	174	2	B34421	male-enhanced antigen	2.35e+03	466	4	21.1	285	2	F71076	hypothetical protein	2.35e+03
394	4	21.1	176	2	A49304	ventral prostate maio	2.35e+03	467	4	21.1	287	4	PC4402	pelA leader/Ig heavy	2.35e+03
395	4	21.1	176	2	JT0816	cystatin-related prot	2.35e+03	468	4	21.1	287	2	JH0332	IgA (Fc) receptor, my	2.35e+03
396	4	21.1	178	2	A38593	transcription factor	2.35e+03	469	4	21.1	288	2	I38485	BCR1123 - human (fra	2.35e+03
397	4	21.1	181	2	A71145	hypothetical protein	2.35e+03	470	4	21.1	289	2	S7161	dodecenoyl-CoA Delta-	2.35e+03
398	4	21.1	183	2	S65815	RNA-directed DNA poly	2.35e+03	471	4	21.1	291	2	T02986	chlorophyll a/b-bind	2.35e+03
399	4	21.1	183	2	S04743	TPA-induced protein 1	2.35e+03	472	4	21.1	291	2	G1300	probable protein-mech	2.35e+03
400	4	21.1	185	2	A34421	male-enhanced antigen	2.35e+03	473	4	21.1	295	2	S50316	CIN5 protein - yeast	2.35e+03
401	4	21.1	186	2	E71209	probable maf protein	2.35e+03	474	4	21.1	296	2	E71118	hypothetical protein	2.35e+03
402	4	21.1	192	2	T02893	hypothetical protein	2.35e+03	475	4	21.1	298	2	A71425	hypothetical protein	2.35e+03
403	4	21.1	192	2	A71084	hypothetical protein	2.35e+03	476	4	21.1	299	2	I49264	ubiquitinating enzyme	2.35e+03
404	4	21.1	194	2	A38203	proline-rich protein	2.35e+03	477	4	21.1	299	2	C70349	succinyl-diaminopimel	2.35e+03
405	4	21.1	195	2	I52578	ALL-1 protein - human	2.35e+03	478	4	21.1	301	2	S66681	peptidylprolyl isomer	2.35e+03
406	4	21.1	195	2	H71266	hypothetical protein	2.35e+03	479	4	21.1	301	2	G71206	probable tryptophanyl	2.35e+03
407	4	21.1	196	2	A71325	hypothetical protein	2.35e+03	480	4	21.1	306	2	E37471	hypothetical helicase	2.35e+03
408	4	21.1	199	2	S63429	Cis1 protein - yeast	2.35e+03	481	4	21.1	308	2	I48080	coatamer complex epsi	2.35e+03
409	4	21.1	199	2	S05509	photosystem II oxygen	2.35e+03	482	4	21.1	308	2	I46019	cell cycle protein Me	2.35e+03
410	4	21.1	209	2	A39759	photosystem I 18K pro	2.35e+03	483	4	21.1	313	2	F70333	growth factor-inducib	2.35e+03
411	4	21.1	209	2	A36344	cortical cytoskeleton	2.35e+03	484	4	21.1	313	2	C71123	aspergillopepsin I (E	2.35e+03
412	4	21.1	211	2	K71262	conserved hypothetica	2.35e+03	485	4	21.1	314	2	A71157	probable sugar transp	2.35e+03
413	4	21.1	213	1	K14MC	adenylate kinase (EC	2.35e+03	486	4	21.1	315	2	C29624	hypothetical protein	2.35e+03
414	4	21.1	214	2	B34503	small nuclear ribonuc	2.35e+03	487	4	21.1	316	2	S26715	spherulin 2a - slime	2.35e+03
415	4	21.1	217	4	I65366	ALL-1/AF-4 mutant fus	2.35e+03	488	4	21.1	317	2	F70462	cell cycle protein Me	2.35e+03
416	4	21.1	218	2	F70458	probable transcriptio	2.35e+03	489	4	21.1	319	2	A36600	growth factor-inducib	2.35e+03
417	4	21.1	219	2	JC4131	glioma pathogenesis-r	2.35e+03	490	4	21.1	320	2	PN0090	aspergillopepsin I (E	2.35e+03
418	4	21.1	219	2	A47399	hypothetical protein	2.35e+03	491	4	21.1	320	2	H69406	conserved hypothetica	2.35e+03
419	4	21.1	219	2	D70411	hypothetical protein	2.35e+03	492	4	21.1	320	2	JC1255	hypothetical protein	2.35e+03
420	4	21.1	220	2	A46597	acidic calmodulin-bn	2.35e+03	493	4	21.1	322	2	T02281	hypothetical protein	2.35e+03
421	4	21.1	227	2	B70438	hypothetical protein	2.35e+03	494	4	21.1	322	2	C69528	conserved hypothetica	2.35e+03
422	4	21.1	231	2	F70471	conserved hypothetica	2.35e+03	495	4	21.1	324	2	D70371	hypothetical protein	2.35e+03
423	4	21.1	233	2	A69755	two-component respons	2.35e+03	496	4	21.1	325	2	F71179	hypothetical protein	2.35e+03
424	4	21.1	235	2	C71354	probable signal pepti	2.35e+03	497	4	21.1	325	2	A47003	cytokine receptor fam	2.35e+03
425	4	21.1	235	2	JC2322	hypothetical 26.8k pr	2.35e+03	498	4	21.1	325	2	B69786	thiamin-monophosphate	2.35e+03
426	4	21.1	239	2	D70359	conserved hypothetica	2.35e+03	499	4	21.1	326	2	S34427	tristetraproline prot	2.35e+03
427	4	21.1	239	2	G03630	Fc alpha hb - human	2.35e+03	500	4	21.1	326	2	A33755	myb-induced myeloid p	2.35e+03
428	4	21.1	239	2	B40710	CD30 ligand - mouse	2.35e+03	501	4	21.1	327	2	G33282	DNA-binding protein (	2.35e+03
429	4	21.1	240	2	S35794	cell division control	2.35e+03	502	4	21.1	330	2	F71146	hypothetical protein	2.35e+03
430	4	21.1	241	2	G71438	probable Clp proteina	2.35e+03	503	4	21.1	331	2	A69794	hypothetical protein	2.35e+03
431	4	21.1	241	2	S63634	ribosomal protein S3	2.35e+03	504	4	21.1	332	2	C71294	hypothetical protein	2.35e+03
432	4	21.1	242	2	H71297	probable outer membra	2.35e+03	505	4	21.1	332	2	F71147	hypothetical protein	2.35e+03
433	4	21.1	243	2	E70007	sepiapterin reductase	2.35e+03	506	4	21.1	334	4	PRHUR	retrovirus-related pr	2.35e+03
434	4	21.1	243	2	A37982	calcium vector protei	2.35e+03	507	4	21.1	335	2	S58892	signaling lymphocytic	2.35e+03
435	4	21.1	251	2	B70450	beta 1.4 glucosyltran	2.35e+03	508	4	21.1	335	2	T02211	probable glucan endo-	2.35e+03
436	4	21.1	252	2	D69763	ferrichrome ABC trans	2.35e+03	509	4	21.1	335	2	A24785	hypothetical protein	2.35e+03
437	4	21.1	254	2	G71041	hypothetical protein	2.35e+03	510	4	21.1	335	2	B71199	hypothetical protein	2.35e+03
438	4	21.1	256	2	F71326	hypothetical protein	2.35e+03	511	4	21.1	336	2	B71366	probable phosphate ac	2.35e+03
439	4	21.1	256	2	A26720	cAMP-regulated M3R pr	2.35e+03	512	4	21.1	337	3	T02303	VP 10 protein - Color	2.35e+03
440	4	21.1	256	2	F71264	hypothetical protein	2.35e+03	513	4	21.1	338	2	D70308	conserved hypothetica	2.35e+03
441	4	21.1	256	2	S16255	hypothetical protein	2.35e+03	514	4	21.1	338	2	C71603	psuedouridine synthet	2.35e+03
442	4	21.1	256	2	D71176	hypothetical protein	2.35e+03	515	4	21.1	338	2	S45578	ISP1 protein - yeast	2.35e+03
443	4	21.1	257	2	I37532	MHC class II histocom	2.35e+03	516	4	21.1	340	2	S51334	TH15 protein homolog	2.35e+03
444	4	21.1	258	2	S33535	myosin heavy chain an	2.35e+03	517	4	21.1	340	2	S35477	TH15 protein - yeast	2.35e+03
445	4	21.1	259	2	S60617	hypothetical protein	2.35e+03	518	4	21.1	340	2	F71376	probable flagellar mo	2.35e+03
446	4	21.1	261	2	H71144	hypothetical protein	2.35e+03	519	4	21.1	341	2	I57997	hypothetical calcium-	2.35e+03
447	4	21.1	263	2	A23659	spectrin beta chain,	2.35e+03	520	4	21.1	341	2	E71564	probable cationic ami	2.35e+03
448	4	21.1	264	2	S01922	hypothetical protein	2.35e+03	521	4	21.1	341	2	S43252	platelet-activating-f	2.35e+03
449	4	21.1	264	2	JC2508	T-cell receptor alpha	2.35e+03	522	4	21.1	341	2	S63666	platelet-activating f	2.35e+03
450	4	21.1	266	2	F27539	testis-specific, vesp	2.35e+03	523	4	21.1	342	2	A40191	platelet-activating f	2.35e+03
451	4	21.1	266	2	H69468	lysophospholipase hom	2.35e+03	524	4	21.1	342	2	S33697	nm1 protein - Asperg	2.35e+03
452	4	21.1	266	2	F71230	hypothetical protein	2.35e+03	525	4	21.1	342	2	H71258	protein Tp33 precurs	2.35e+03
453	4	21.1	266	2	G71144	hypothetical protein	2.35e+03	526	4	21.1	343	2	T01068	calreticulin - Africa	2.35e+03
454	4	21.1	267	2	G70455	hypothetical protein	2.35e+03	527	4	21.1	343	2	S17677	NADH dehydrogenase (u	2.35e+03
455	4	21.1	271	2	A49586	major histocompatibil	2.35e+03	528	4	21.1	343	2	A70386	hypothetical protein	2.35e+03
456	4	21.1	271	2	D71230	hypothetical protein	2.35e+03	529	4	21.1	345	2	H71358	conserved hypothetica	2.35e+03
457	4	21.1	273	2	G01418	cytokine receptor fam	2.35e+03	530	4	21.1	345	2	C71046	hypothetical protein	2.35e+03
458	4	21.1	273	2	D70428	hypothetical protein	2.35e+03	531	4	21.1	345	2	C71237	probable transport-A	2.35e+03
459	4	21.1	275	2	S08464	T-cell alloantigen RT	2.35e+03	532	4	21.1	346	2	C71042	probable threonine-S	2.35e+03
460	4	21.1	275	2	E71269	probable riboflavin k	2.35e+03	533	4	21.1	349	2	A53340	interferon regulatory	2.35e+03
461	4	21.1	275	2	A34866	T-cell surface protei	2.35e+03	534	4	21.1	349	2	T00526	probable proline-rich	2.35e+03



681	4	21.1	2	C70485	poly A polymerase - A	2.35e+03	754	4	21.1	2	G71306	hypothetical protein	2.35e+03	
682	4	21.1	3	JEO361	cytochromes P450, CYP	2.35e+03	755	4	21.1	616	2	A55796	ecarin precursor - sa	2.35e+03
683	4	21.1	515	A43289	CD44L protein - human	2.35e+03	756	4	21.1	616	2	T00894	hypothetical protein	2.35e+03
684	4	21.1	517	B71260	hypothetical protein	2.35e+03	757	4	21.1	617	2	S52797	muf1 protein - human	2.35e+03
685	4	21.1	2	A45982	calcium channel beta	2.35e+03	758	4	21.1	619	2	A56519	CDCL6 protein - human	2.35e+03
686	4	21.1	519	B45697	immediate-early prote	2.35e+03	759	4	21.1	621	2	138467	low density lipoprote	2.35e+03
687	4	21.1	521	F151693	xpologomb - African c	2.35e+03	760	4	21.1	623	2	146382	zpi precursor - mouse	2.35e+03
688	4	21.1	521	F71009	probable sodium-depen	2.35e+03	761	4	21.1	626	2	149100	mscd6 precursor - mou	2.35e+03
689	4	21.1	522	165767	L-type voltage-gated	2.35e+03	762	4	21.1	630	2	S29796	hypothetical protein	2.35e+03
690	4	21.1	523	S42515	recombination-activat	2.35e+03	763	4	21.1	632	2	T00679	hypothetical protein	2.35e+03
691	4	21.1	523	S42514	recombination-activat	2.35e+03	764	4	21.1	632	2	T00325	hypothetical protein	2.35e+03
692	4	21.1	524	A41347	calcium channel prote	2.35e+03	765	4	21.1	633	2	T02673	heterogeneous nuclear	2.35e+03
693	4	21.1	527	S42512	RAG-2 protein - rabbi	2.35e+03	766	4	21.1	635	2	A45266	MPL-P protein precurs	2.35e+03
694	4	21.1	528	1K1BYG	galactokinase (EC 2.7	2.35e+03	767	4	21.1	644	2	S35109	cell division protein	2.35e+03
695	4	21.1	528	A41718	polypyrimidine tract-	2.35e+03	768	4	21.1	645	2	G01205	tyl protein - human	2.35e+03
696	4	21.1	530	A36340	ELAM-1 ligand fucosyl	2.35e+03	769	4	21.1	648	2	A57284	spermatid perinuclear	2.35e+03
697	4	21.1	530	S13552	polypyrimidine tract-	2.35e+03	770	4	21.1	649	2	YN0809	drebrin E (clone gpbh	2.35e+03
698	4	21.1	531	S23015	polypyrimidine tract-	2.35e+03	771	4	21.1	653	2	B71157	hypothetical protein	2.35e+03
699	4	21.1	533	J03034	developmental control	2.35e+03	772	4	21.1	656	2	S73313	hypothetical protein	2.35e+03
700	4	21.1	533	A37180	chromogranin/secretog	2.35e+03	773	4	21.1	657	2	A53545	protein p84 - human	2.35e+03
701	4	21.1	535	JC5762	cytokine-inducible SH	2.35e+03	774	4	21.1	660	3	B54942	acyl-CoA oxidase (EC	2.35e+03
702	4	21.1	540	A47022	chitinase (EC 3.2.1.1	2.35e+03	775	4	21.1	660	3	T02768	glycoprotein G homolo	2.35e+03
703	4	21.1	542	B47022	chitinase (EC 3.2.1.1	2.35e+03	776	4	21.1	661	2	152603	MPS1 protein - mouse	2.35e+03
704	4	21.1	543	A41211	early growth response	2.35e+03	777	4	21.1	666	2	S61461	p83/100 protein - Bor	2.35e+03
705	4	21.1	547	T02274	hypothetical protein	2.35e+03	778	4	21.1	669	2	E71610	hypothetical protein	2.35e+03
706	4	21.1	550	S23016	polypyrimidine tract-	2.35e+03	779	4	21.1	671	2	JC5496	prox 1 protein 671 -	2.35e+03
707	4	21.1	556	S36629	polypyrimidine tract-	2.35e+03	780	4	21.1	672	2	A41075	inositol-1,4,5-trisph	2.35e+03
708	4	21.1	557	S26294	polypyrimidine tract-	2.35e+03	781	4	21.1	673	2	S41053	1D-myo-inositol-trisp	2.35e+03
709	4	21.1	557	JEO346	high-affinity carntin	2.35e+03	782	4	21.1	676	2	T01084	probable protein dest	2.35e+03
710	4	21.1	559	F71041	probable DNA ligase -	2.35e+03	783	4	21.1	677	2	S57599	FUS2 protein - yeast	2.35e+03
711	4	21.1	559	I19444	SH3 binding protein -	2.35e+03	784	4	21.1	677	2	B71306	hypothetical protein	2.35e+03
712	4	21.1	559	B56731	chromatin assembly fa	2.35e+03	785	4	21.1	679	2	T00636	hypothetical protein	2.35e+03
713	4	21.1	560	H70466	serine proteinase (EC	2.35e+03	786	4	21.1	680	2	A43800	nuclear autoantigenic	2.35e+03
714	4	21.1	564	B43776	drebrin E1 - chicken	2.35e+03	787	4	21.1	683	2	JC5393	zinc finger protein K	2.35e+03
715	4	21.1	564	B70409	hypothetical protein	2.35e+03	788	4	21.1	684	2	D71683	proteinase II (ptrB)	2.35e+03
716	4	21.1	565	1HMIVT3	hemagglutinin precurs	2.35e+03	789	4	21.1	685	2	JC5392	zinc finger protein K	2.35e+03
717	4	21.1	568	I19031	poly(A) polymerase -	2.35e+03	790	4	21.1	686	2	JC5708	villin-like protein -	2.35e+03
718	4	21.1	568	S38084	hypothetical protein -	2.35e+03	791	4	21.1	690	2	H71237	hypothetical protein	2.35e+03
719	4	21.1	570	B44282	retrovirus-related en	2.35e+03	792	4	21.1	692	2	A45991	protransglutaminase E	2.35e+03
720	4	21.1	572	S61650	hypothetical protein	2.35e+03	793	4	21.1	703	2	S45686	glycogen synthase - h	2.35e+03
721	4	21.1	576	A48157	renal transcription f	2.35e+03	794	4	21.1	703	2	A35362	UDPglucose-glycogen	2.35e+03
722	4	21.1	579	JH0820	160K golgi antigen -	2.35e+03	795	4	21.1	704	2	T02558	hypothetical protein	2.35e+03
723	4	21.1	579	S63069	hypothetical protein	2.35e+03	796	4	21.1	706	2	H70425	conserved hypothetical	2.35e+03
724	4	21.1	579	S46210	precursor - common to	2.35e+03	797	4	21.1	707	2	S60588	drebrin A - rat	2.35e+03
725	4	21.1	579	T02574	hypothetical protein	2.35e+03	798	4	21.1	707	2	F71290	probable flagellar bi	2.35e+03
726	4	21.1	584	S45266	hypothetical protein	2.35e+03	799	4	21.1	720	2	G70320	aldehyde dehydrogenas	2.35e+03
727	4	21.1	584	S40013	hypothetical protein	2.35e+03	800	4	21.1	725	2	A45033	myelin transcription	2.35e+03
728	4	21.1	585	G02165	human homolog of Dros	2.35e+03	801	4	21.1	735	2	I48101	ADAM 6 protein precur	2.35e+03
729	4	21.1	587	JEO268	myosin - mouse	2.35e+03	802	4	21.1	735	2	A33369	UDPglucose-glycogen	2.35e+03
730	4	21.1	588	T03740	glucose-6-phosphate 1	2.35e+03	803	4	21.1	736	2	JC5495	prox 1 protein - chic	2.35e+03
731	4	21.1	590	S63193	hypothetical protein	2.35e+03	804	4	21.1	736	2	G01532	acidic 82 kDa protein	2.35e+03
732	4	21.1	591	S65088	finger protein XPO 6	2.35e+03	805	4	21.1	737	2	JE0269	Proxi protein - mouse	2.35e+03
733	4	21.1	592	C70302	conserved hypothetical	2.35e+03	806	4	21.1	737	2	A32156	UDPglucose-glycogen	2.35e+03
734	4	21.1	592	F71423	hypothetical protein	2.35e+03	807	4	21.1	739	2	S15727	cellulase (EC 3.2.1.4	2.35e+03
735	4	21.1	593	T03244	glucose-6-phosphate 1	2.35e+03	808	4	21.1	739	2	I40715	malate synthase (EC 4	2.35e+03
736	4	21.1	593	I51213	drebrin - chicken (fr	2.35e+03	809	4	21.1	740	2	F71369	conserved hypothetical	2.35e+03
737	4	21.1	594	I49127	intracellular protein	2.35e+03	810	4	21.1	745	2	I57665	gene LRPR1 protein -	2.35e+03
738	4	21.1	596	C44461	voltage-dependent cal	2.35e+03	811	4	21.1	747	2	F69332	heterodisulfide reduc	2.35e+03
739	4	21.1	597	T18304	calcium channel prote	2.35e+03	812	4	21.1	750	2	A56881	prostate-specific mem	2.35e+03
740	4	21.1	598	I52859	L-type voltage-gated	2.35e+03	813	4	21.1	754	2	A53679	endothelin converting	2.35e+03
741	4	21.1	598	H71362	probable tpr protein	2.35e+03	814	4	21.1	754	2	S47268	endothelin-converting	2.35e+03
742	4	21.1	598	D71365	probable tpr protein	2.35e+03	815	4	21.1	758	2	JC2521	endothelin converting	2.35e+03
743	4	21.1	600	S58681	selenocysteine tRNA a	2.35e+03	816	4	21.1	758	2	A54667	endothelin converting	2.35e+03
744	4	21.1	602	T03202	VP10' protein - Color	2.35e+03	817	4	21.1	761	2	G70393	conserved hypothetical	2.35e+03
745	4	21.1	604	S24760	hydroxymethylglutaryl	2.35e+03	818	4	21.1	761	2	A47563	glucose-6-phosphate 1	2.35e+03
746	4	21.1	605	T0A0030	hypothetical protein	2.35e+03	819	4	21.1	768	2	A42755	p-selectin precursor	2.35e+03
747	4	21.1	606	H71065	hypothetical protein	2.35e+03	820	4	21.1	768	2	I53821	p-selectin - rat	2.35e+03
748	4	21.1	607	A43776	drebrin E2 - Chicken	2.35e+03	821	4	21.1	770	2	JC4136	endothelin converting	2.35e+03
749	4	21.1	608	S72177	dextranase (EC 3.2.1.	2.35e+03	822	4	21.1	770	2	G02228	DOC-2 - human	2.35e+03
750	4	21.1	609	T07301	probable tpr protein	2.35e+03	823	4	21.1	771	2	S40956	membrane protein SKN1	2.35e+03
751	4	21.1	610	A35046	endothelial leukocyte	2.35e+03	824	4	21.1	771	2	B38252	granulocyte colony-st	2.35e+03
752	4	21.1	612	A32606	endothelial leukocyte	2.35e+03	825	4	21.1	772	2	A41860	tetracycline resistan	2.35e+03
753	4	21.1	613	F69424	conserved hypothetical	2.35e+03	826	4	21.1	774	2	JC2299	cell surface glycopro	2.35e+03

827	4	21.1	775	2	I49237	2.35e+03	900	4	21.1	1074	2	JC5928	2.35e+03	semaphorin F precursor	2.35e+03
828	4	21.1	776	2	T02702	2.35e+03	901	4	21.1	1083	2	S54293	2.35e+03	regulator protein p12	2.35e+03
829	4	21.1	778	2	E71263	2.35e+03	902	4	21.1	1085	2	S40476	2.35e+03	Ca(2+)-sensing recept	2.35e+03
830	4	21.1	780	2	S39110	2.35e+03	903	4	21.1	1116	2	B70476	2.35e+03	hypothetical protein	2.35e+03
831	4	21.1	782	2	A31491	2.35e+03	904	4	21.1	1122	1	NCECXV	2.35e+03	exodeoxyribonuclease	2.35e+03
832	4	21.1	783	2	JH0329	2.35e+03	905	4	21.1	1127	2	S47445	2.35e+03	MDM1 protein - yeast	2.35e+03
833	4	21.1	786	2	I49274	2.35e+03	906	4	21.1	1132	2	T00259	2.35e+03	hypothetical protein	2.35e+03
834	4	21.1	786	2	S54773	2.35e+03	907	4	21.1	1132	2	A35089	2.35e+03	myosin-binding protein	2.35e+03
835	4	21.1	789	2	T01321	2.35e+03	908	4	21.1	1150	2	B47114	2.35e+03	phosphoprotein phosph	2.35e+03
836	4	21.1	790	2	A35797	2.35e+03	909	4	21.1	1150	2	A41641	2.35e+03	mannosyl-oligosacchar	2.35e+03
837	4	21.1	792	2	A70476	2.35e+03	910	4	21.1	1151	2	H71347	2.35e+03	hypothetical protein	2.35e+03
838	4	21.1	793	2	S54772	2.35e+03	911	4	21.1	1170	2	A57650	2.35e+03	repair protein XPG -	2.35e+03
839	4	21.1	794	2	G02317	2.35e+03	912	4	21.1	1171	2	T00380	2.35e+03	KIAA0637 protein - hu	2.35e+03
840	4	21.1	794	2	S55527	2.35e+03	913	4	21.1	1174	1	HJBYDH	2.35e+03	helicase (EC 3.6.1.-)	2.35e+03
841	4	21.1	797	2	S44546	2.35e+03	914	4	21.1	1179	2	F71190	2.35e+03	probable chromosome a	2.35e+03
842	4	21.1	799	2	T00331	2.35e+03	915	4	21.1	1188	2	CJ4889	2.35e+03	phosphatidylinositol	2.35e+03
843	4	21.1	803	2	S26823	2.35e+03	916	4	21.1	1188	2	C71231	2.35e+03	hypothetical protein	2.35e+03
844	4	21.1	807	2	A34581	2.35e+03	917	4	21.1	1189	2	JC6118	2.35e+03	SH2-containing inosit	2.35e+03
845	4	21.1	809	2	A34404	2.35e+03	918	4	21.1	1196	2	S46430	2.35e+03	botulinum neurotoxin-	2.35e+03
846	4	21.1	811	3	T00036	2.35e+03	919	4	21.1	1216	2	A55620	2.35e+03	apical endosomal prot	2.35e+03
847	4	21.1	813	2	JC6569	2.35e+03	920	4	21.1	1217	2	T00270	2.35e+03	hypothetical protein	2.35e+03
848	4	21.1	823	2	H71300	2.35e+03	921	4	21.1	1224	2	T00059	2.35e+03	hypothetical protein	2.35e+03
849	4	21.1	824	2	I49564	2.35e+03	922	4	21.1	1230	2	S53974	2.35e+03	hypothetical protein	2.35e+03
850	4	21.1	825	2	I46078	2.35e+03	923	4	21.1	1232	2	S40766	2.35e+03	hypothetical protein	2.35e+03
851	4	21.1	830	2	B44439	2.35e+03	924	4	21.1	1232	2	A54748	2.35e+03	neuronal apoptosis in	2.35e+03
852	4	21.1	837	2	A57542	2.35e+03	925	4	21.1	1240	2	T03097	2.35e+03	CDO protein - human	2.35e+03
853	4	21.1	839	2	I50590	2.35e+03	926	4	21.1	1251	2	A56677	2.35e+03	neuronal cell cycle w	2.35e+03
854	4	21.1	840	2	T02164	2.35e+03	927	4	21.1	1257	2	S28764	2.35e+03	neurocan - rat	2.35e+03
855	4	21.1	845	2	A40016	2.35e+03	928	4	21.1	1282	2	FJ0120	2.35e+03	glycoprotein A - mous	2.35e+03
856	4	21.1	845	2	H71317	2.35e+03	929	4	21.1	1326	2	B56395	2.35e+03	secretory phospholipa	2.35e+03
857	4	21.1	856	2	A44439	2.35e+03	930	4	21.1	1347	2	T02214	2.35e+03	ubiquitous TPR motif	2.35e+03
858	4	21.1	858	1	ZPECPA	2.35e+03	931	4	21.1	1352	2	G71051	2.35e+03	probable ATP-dependen	2.35e+03
859	4	21.1	863	2	S38252	2.35e+03	932	4	21.1	1365	2	T00833	2.35e+03	probable RNA-directed	2.35e+03
860	4	21.1	869	2	A55384	2.35e+03	933	4	21.1	1384	3	T02748	2.35e+03	hypothetical protein	2.35e+03
861	4	21.1	877	2	I50591	2.35e+03	934	4	21.1	1390	2	A45455	2.35e+03	nucleoporin 155 - rat	2.35e+03
862	4	21.1	888	1	RDHUE	2.35e+03	935	4	21.1	1401	2	T02255	2.35e+03	probable ubiquitously	2.35e+03
863	4	21.1	889	2	T02240	2.35e+03	936	4	21.1	1418	2	S40764	2.35e+03	hypothetical protein	2.35e+03
864	4	21.1	900	2	S70630	2.35e+03	937	4	21.1	1421	2	T00333	2.35e+03	hypothetical protein	2.35e+03
865	4	21.1	901	3	T03726	2.35e+03	938	4	21.1	1427	2	I51669	2.35e+03	tumor suppressor - Af	2.35e+03
866	4	21.1	902	2	A60560	2.35e+03	939	4	21.1	1429	2	JM0609	2.35e+03	nitric-oxide synthase	2.35e+03
867	4	21.1	903	2	T00074	2.35e+03	940	4	21.1	1429	2	S16233	2.35e+03	nitric-oxide synthase	2.35e+03
868	4	21.1	908	2	B69435	2.35e+03	941	4	21.1	1433	2	G01946	2.35e+03	nitric-oxide synthase	2.35e+03
869	4	21.1	911	2	A56465	2.35e+03	942	4	21.1	1434	2	G71232	2.35e+03	hypothetical protein	2.35e+03
870	4	21.1	921	2	A54139	2.35e+03	943	4	21.1	1441	2	T00335	2.35e+03	hypothetical protein	2.35e+03
871	4	21.1	921	2	G02326	2.35e+03	944	4	21.1	1447	2	A54100	2.35e+03	tumor suppressor prot	2.35e+03
872	4	21.1	935	2	S43730	2.35e+03	945	4	21.1	1462	2	T00345	2.35e+03	hypothetical protein	2.35e+03
873	4	21.1	940	2	S44345	2.35e+03	946	4	21.1	1477	2	T00957	2.35e+03	hypothetical protein	2.35e+03
874	4	21.1	941	2	B53335	2.35e+03	947	4	21.1	1481	2	A47341	2.35e+03	amylopullulanase prec	2.35e+03
875	4	21.1	943	2	A42681	2.35e+03	948	4	21.1	1493	2	A44224	2.35e+03	DNA repair helicase E	2.35e+03
876	4	21.1	947	2	I49635	2.35e+03	949	4	21.1	1495	2	S22610	2.35e+03	DNA (cytosine-5-)-met	2.35e+03
877	4	21.1	957	2	A53789	2.35e+03	950	4	21.1	1516	2	E71619	2.35e+03	RAD2 endonuclease PFB	2.35e+03
878	4	21.1	958	2	S32435	2.35e+03	951	4	21.1	1525	2	T01661	2.35e+03	probable DNA (cytosin	2.35e+03
879	4	21.1	961	2	JC2200	2.35e+03	952	4	21.1	1571	2	T00062	2.35e+03	hypothetical protein	2.35e+03
880	4	21.1	966	1	PIJBVB	2.35e+03	953	4	21.1	1572	2	T00376	2.35e+03	KIAA0645 protein - hu	2.35e+03
881	4	21.1	970	2	I48037	2.35e+03	954	4	21.1	1584	2	T00026	2.35e+03	brain-specific angio	2.35e+03
882	4	21.1	970	2	S27114	2.35e+03	955	4	21.1	1616	2	S62429	2.35e+03	hypothetical protein	2.35e+03
883	4	21.1	970	2	A56767	2.35e+03	956	4	21.1	1627	2	S65464	2.35e+03	pregnancy-associated	2.35e+03
884	4	21.1	970	2	S36417	2.35e+03	957	4	21.1	1630	2	S64403	2.35e+03	ESPI protein - yeast	2.35e+03
885	4	21.1	971	2	S28833	2.35e+03	958	4	21.1	1643	2	D71630	2.35e+03	outer membrane protei	2.35e+03
886	4	21.1	973	2	S32815	2.35e+03	959	4	21.1	1722	2	I78879	2.35e+03	retinoblastoma bindin	2.35e+03
887	4	21.1	981	2	S55132	2.35e+03	960	4	21.1	1733	2	S27939	2.35e+03	tensin - chicken	2.35e+03
888	4	21.1	988	1	DJVF2FP	2.35e+03	961	4	21.1	1737	2	A37491	2.35e+03	hypothetical helicase	2.35e+03
889	4	21.1	997	2	JC6067	2.35e+03	962	4	21.1	1744	2	A54970	2.35e+03	tensin, cardiac muscl	2.35e+03
890	4	21.1	1002	2	C70319	2.35e+03	963	4	21.1	1744	2	JH0720	2.35e+03	tenascin - African cla	2.35e+03
891	4	21.1	1014	2	S48235	2.35e+03	964	4	21.1	1756	2	T02599	2.35e+03	hypothetical protein	2.35e+03
892	4	21.1	1016	2	T00375	2.35e+03	965	4	21.1	1792	2	A57075	2.35e+03	tensin - chicken (fra	2.35e+03
893	4	21.1	1038	2	JC5497	2.35e+03	966	4	21.1	1805	2	A34736	2.35e+03	nestin - rat	2.35e+03
894	4	21.1	1047	1	OYHU6R	2.35e+03	967	4	21.1	1811	3	T00035	2.35e+03	nonstructural polypro	2.35e+03
895	4	21.1	1052	2	T00067	2.35e+03	968	4	21.1	1849	2	T00415	2.35e+03	hypothetical protein	2.35e+03
896	4	21.1	1056	2	S56039	2.35e+03	969	4	21.1	1870	2	S37671	2.35e+03	bat2 protein - human	2.35e+03
897	4	21.1	1061	1	OYHUAR	2.35e+03	970	4	21.1	1872	2	S36152	2.35e+03	bat2 protein - human	2.35e+03
898	4	21.1	1061	2	A57620	2.35e+03	971	4	21.1	1882	2	T00069	2.35e+03	hypothetical protein	2.35e+03
899	4	21.1	1066	2	S50237	2.35e+03	972	4	21.1	1886	2	S04921	2.35e+03	nuclear pore protein	2.35e+03

```
973 4 21.1 1916 2 S46157 2.35e+03
974 4 21.1 1926 2 S01169 2.35e+03
975 4 21.1 1938 2 I37453 2.35e+03
976 4 21.1 1944 2 A55117 2.35e+03
977 4 21.1 1956 2 T00051 2.35e+03
978 4 21.1 2024 2 A54103 2.35e+03
979 4 21.1 2035 2 A40718 2.35e+03
980 4 21.1 2142 2 B35098 2.35e+03
981 4 21.1 2148 2 A56081 2.35e+03
982 4 21.1 2216 2 S78398 2.35e+03
983 4 21.1 2500 2 G71009 2.35e+03
984 4 21.1 2649 2 A40937 2.35e+03
985 4 21.1 2818 2 B55282 2.35e+03
986 4 21.1 2820 2 J53196 2.35e+03
987 4 21.1 2871 2 A55624 2.35e+03
988 4 21.1 2871 2 A55567 2.35e+03
989 4 21.1 2897 2 B48666 2.35e+03
990 4 21.1 3002 2 A47221 2.35e+03
991 4 21.1 3224 2 S58884 2.35e+03
992 4 21.1 3256 2 A48666 2.35e+03
993 4 21.1 3866 2 B48205 2.35e+03
994 4 21.1 3869 2 A48205 2.35e+03
995 4 21.1 3910 2 A44264 2.35e+03
996 4 21.1 3968 2 A44265 2.35e+03
997 4 21.1 4957 2 T03455 2.35e+03
998 4 21.1 5262 2 T03454 2.35e+03
999 4 21.1 7962 2 I38346 2.35e+03
1000 4 21.1 15281 2 S41309 2.35e+03
```

## ALIGNMENTS

```
RESULT 1
ENTRY JG0168 #type complete
TITLE gob-5 protein - Mouse
ORGANISM gob-5 protein Mus musculus #common_name house mouse
DATE 18-Mar-1999 #sequence_revision 18-Mar-1999 #text_change
18-Mar-1999

ACCESSIONS JG0168
REFERENCE JG0168
#authors Komiya, T.; Tanigawa, Y.; Hirohashi, S.
#journal Biochem. Biophys. Res. Commun. (1999) 255:347-351
#title Cloning and identification of the gene gob-5, which is
expressed in intestinal goblet cells in mice.
#accession JG0168
#status preliminary
#residues 1-913 #label KOM
#cross-references DDBJ:AB016592
SUMMARY #length 913 #molecular-weight 100070 #checksum 8755

Query Match 52.6%; Score 10; DB 3; Length 913;
Best Local Similarity 100.0%; Pred. No. 8.88e-08;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 169 YNNDEKFYLS 178
|||||
QY 1 YNNDEKFYLS 10

RESULT 2
ENTRY D71614 #type complete
TITLE hypothetical protein PFB0460c - malaria parasite (Plasmodium
falciparum)
ORGANISM #formal_name Plasmodium falciparum
DATE 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change
13-Nov-1998
D71614
A71600
#accessions Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.;
#authors Aravind, L.; Koonin, E.V.; Shalton, S.; Mason, T.; Yu, K.;
Fuji, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.;
Lai, Z.; Schwartz, D.C.; Pertea, M.; Salzberg, S.; Zhou,
L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
```

```
#journal Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.L.
#title Science (1998) 282:1126-1132
#accession Chromosome 2 sequence of the human malaria parasite
#status Plasmodium falciparum.
D71614
#molecule_type DNA
#residues 1-2573 #label GAR
#cross-references GB:AE001396; GB:AE001362; NID:g3845188; PID:g3845190;
#experimental_source clone 3d7
TIGR:PFB0460c
GENETICS
#gene PFB0460c
SUMMARY #length 2573 #molecular-weight 308142 #checksum 4299

Query Match 36.8%; Score 7; DB 2; Length 2573;
Best Local Similarity 100.0%; Pred. No. 5.35e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 2212 YNNDEKF 2218
|||||
QY 1 YNNDEKF 7

RESULT 3
ENTRY G69121 #type complete
TITLE hypothetical protein MTH1908 - Methanobacterium
thermoautotrophicum (strain Delta H)
ORGANISM #formal_name Methanobacterium thermoautotrophicum
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
05-Jun-1998

ACCESSIONS G69121
REFERENCE A69000
#authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.;
Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso,
A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;
McDougall, S.; Shimer, G.; Goyal, A.; Pietrovskiy, S.;
Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
J.; Reeve, J.N.
#journal J. Bacteriol. (1997) 179:7135-7155
#title Complete genome sequence of Methanobacterium
thermoautotrophicum Delta H: functional analysis and
comparative genomics.
```

```
#cross-references MUID:98037514
#accession G69121
#status preliminary; nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-97 #label MTH
#cross-references GB:AE000942; GB:AE000666; NID:g2623039; PID:g2623042
#experimental_source strain Delta H
GENETICS
#gene MTH1908
SUMMARY #length 97 #molecular-weight 10432 #checksum 8423

Query Match 31.6%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.84e-00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 23 FYLSNG 28
|||||
QY 7 FYLSNG 12

RESULT 4
ENTRY S37812 #type complete
TITLE hypothetical protein YKL002w - yeast (Saccharomyces
cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
```

```

DATE          03-May-1994 #sequence_revision 03-May-1994 #text_change
ACCESSIONS    S37812; S37813
REFERENCE      Duesterhoeft, A.; Moestl, D.; Poehlmann, R.; Philippssen, P.
#authors      submitted to the Protein Sequence Database, March 1994
#accession    S37812
#molecule_type DNA
##residues    1-151 #label DUE
##cross-references EMBL:Z28002; NID:g485988; PID:g485989; MIPS:YKL002W
##experimental_source strain S288C
REFERENCE      S37813
#authors      Boyer, J.; Pascolo, S.; Richard, G.F.; Ghazvini, M.;
#              Colleaax, L.; Thierry, A.; Monnier, A.L.; Dujon, B.
#submission    submitted to the Protein Sequence Database, March 1994
#accession    S37813
#molecule_type DNA
##residues    1-151 #label BOY
##cross-references EMBL:Z28002; NID:g485988; PID:g485989; MIPS:YKL002W
##experimental_source strain S288C
GENETICS
#map_position 1L1
SUMMARY        #length 151 #molecular-weight 16679 #checksum 5962
Query Match    31.6%; Score 6; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 11 RIQAVR 16
| | | | |
QY 13 RIQAVR 18

RESULT 5
ENTRY   S73534 #type complete
TITLE   hypothetical protein C12orf181o - Mycoplasma pneumoniae
        (ATCC 29342) (SGC3)
ORGANISM #formal_name Mycoplasma pneumoniae
#variety ATCC 29342
DATE     27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change
        17-Jul-1998
ACCESSIONS S73534
REFERENCE   S73327
#authors    Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li,
#              B.C.; Herrmann, R.
#journal    Nucleic Acids Res. (1996) 24:4420-4449
#title      Complete sequence analysis of the genome of the bacterium
#              Mycoplasma pneumoniae.
#cross-references MUID:97105885
#accession  S73534
#status     preliminary; nucleic acid sequence not shown;
#              translation not shown
#molecule_type DNA
##residues  1-181 #label HIM
##cross-references EMBL:AE000021; GB:U00089; NID:g1673871; PID:g1673873
#note        the nucleotide sequence was submitted to the EMBL Data
              Library, November 1996
GENETICS
#genetic_code SGC3
SUMMARY        #length 181 #molecular-weight 21358 #checksum 5728
Query Match    31.6%; Score 6; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 125 SNGRIQ 130
| | | | |
QY 10 SNGRIQ 15

RESULT 6
ENTRY   S58329 #type complete
TITLE   HSH49 protein - yeast (Saccharomyces cerevisiae)

```

```

ALTERNATE_NAMES protein O6142; protein YOR319W
ORGANISM          #formal_name Saccharomyces cerevisiae
DATE              13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change
ACCESSIONS        S58329; S67225; S71999
REFERENCE          S58318
#authors          Pearson, B.M.; Hernando, Y.; Wolf, S.S.; Kalogeropoulos, A.;
#              Schweizer, M.
#submission        submitted to the EMBL Data Library, August 1995
#accession         S58329
#molecule_type    DNA
##residues         1-213 #label PEA
##cross-references EMBL:X90565; NID:g940836; PID:g940850
REFERENCE          S67213
#authors          Pearson, B.M.; Hernando, Y.; Kalogeropoulos, A.; Schweizer,
#              M.
#submission        submitted to the Protein Sequence Database, July 1996
#accession         S67225
#molecule_type    DNA
##residues         1-213 #label PEW
##cross-references EMBL:Z75227; NID:g1420699; PID:e252149; PID:g1420700;
#              MIPS:YOR319W
#experimental_source strain S288C
REFERENCE          S71986
#authors          Pearson, B.M.; Hernando, Y.; Payne, J.; Wolf, S.S.;
#              Kalogeropoulos, A.; Schweizer, M.
#journal           Yeast (1996) 12:1021-1031
#title            Sequencing of a 35.71 kb DNA segment on the right arm of
#              yeast chromosome XV reveals regions of similarity to
#              chromosomes I and XIII.
#accession         S71999
#status            nucleic acid sequence not shown; translation not shown
#molecule_type    DNA
##residues         1-213 #label PEF
#note             the nucleotide sequence was submitted to the EMBL Data
              Library, August 1995
GENETICS
#gene           SGD:HS49
#map_position    15R
#cross-references SGD:S0005846; MIPS:YOR319W
FUNCTION          essential splicing factor
CLASSIFICATION    #superfamily ribonucleoprotein repeat homology
KEYWORDS          nucleus
FEATURE           10-78
109-175           #domain ribonucleoprotein repeat homology #label RRM1\
#domain ribonucleoprotein repeat homology #label RRM2
SUMMARY          #length 213 #molecular-weight 24503 #checksum 5177
Query Match      31.6%; Score 6; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 141 FYLSNG 146
| | | | |
QY 7 FYLSNG 12

RESULT 7
ENTRY   S75995 #type complete
TITLE   hypothetical protein - Synecocystis sp. (strain PCC 6803)
ORGANISM #formal_name Synecocystis sp.
#variety PCC 6803
DATE     25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change
        21-Aug-1998
ACCESSIONS S75995
REFERENCE   S74322
#authors    Kaneo, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.;
#              Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugita, M.;
#              Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.;
#              Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimo,
#              S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;

```

```
#journal      Yasuda, M.; Tabata, S.  
#description  DNA Res. (1996) 3:109-136  
#title        Sequence analysis of the genome of the unicellular  
              cyanobacterium Synechocystis sp. PCC6803. II. Sequence  
              determination of the entire genome and assignment of  
              potential protein-coding regions.  
#cross-references EMBL:D64006; GB:AB001339; NID:g1001291; PID:d1011493;  
#accession      S75995  
#status        preliminary  
#molecule_type DNA  
#residues       1-314 #label KAN  
#cross-references EMBL:D64006; GB:AB001339; NID:g1001291; PID:d1011493;  
#note          the nucleotide sequence was submitted to the EMBL Data  
              Library, June 1996  
SUMMARY       #length 314 #molecular-weight 34632 #checksum 9177  
  
Query Match   31.6%; Score 6; DB 2; Length 314;  
Best Local Similarity 100.0%; Pred. No. 2.84e+00;  
Matches       6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 144 GRIQAV 149  
QY 12 GRIQAV 17  
      |||||  
RESULT 8  
ENTRY   #type fragment  
TITLE   hypothetical protein MCL00 - Mycoplasma capricolum (SGC3)  
ORGANISM #formal_name Mycoplasma capricolum  
DATE     09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change  
21-Nov-1997  
ACCESSIONS S7783  
REFERENCE   Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander,  
#authors    C.; Dolan, M.; Gilbert, W.; Gillevet, P.M.  
#journal    Mol. Microbiol. (1995) 16:955-967  
#title      Exploring the Mycoplasma capricolum genome: a minimal cell  
              reveals its physiology.  
#cross-references MUID:96059641  
#accession  S7783  
#status     nucleic acid sequence not shown; translation not shown  
#molecule_type DNA  
#residues   1-316 #label BOR  
#cross-references EMBL:D33074; NID:g516148; PID:g530423  
#experimental_source ATCC 27343  
#note       the nucleotide sequence was submitted to the EMBL Data  
              Library, July 1994  
GENETICS  
#genetic_code SGC3  
SUMMARY       #length 316 #checksum 1795  
  
Query Match   31.6%; Score 6; DB 2; Length 316;  
Best Local Similarity 100.0%; Pred. No. 2.84e+00;  
Matches       6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 173 NDEKFY 178  
QY 3 NDEKFY 8  
      |||||  
RESULT 9  
ENTRY   #type complete  
TITLE   xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-) XTR-2 -  
Arabidopsis thaliana  
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear  
              cross  
DATE     28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change  
30-Oct-1998  
ACCESSIONS S71224  
REFERENCE   Xu, W.; Campbell, P.; Vargheese, A.K.; Braam, J.  
#authors
```

```
#submission    submitted to the EMBL Data Library, December 1995  
#description    The Arabidopsis XTR-related gene family: Environmental and  
              hormonal regulation of expression.  
#accession      S71224  
#molecule_type mRNA  
#residues       1-332 #label XUV  
#cross-references EMBL:U43487; NID:g1244755; PID:g1244756  
GENETICS  
#gene          XTR2  
CLASSIFICATION #superfamily endoxylglucan transferase  
KEYWORDS       glycosidase; hydrolase  
SUMMARY       #length 332 #molecular-weight 38252 #checksum 3110  
  
Query Match   31.6%; Score 6; DB 2; Length 332;  
Best Local Similarity 100.0%; Pred. No. 2.84e+00;  
Matches       6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 94 FYLSNG 99  
QY 7 FYLSNG 12  
      |||||  
RESULT 10  
ENTRY   #type complete  
TITLE   naringenin 3-dioxygenase (EC 1.14.11.9) - China aster  
ORGANISM #formal_name Callistephus chinensis #common_name China aster  
DATE     02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change  
17-Mar-1999  
ACCESSIONS S38336; S32147  
REFERENCE   Britsch, L.; Dedio, J.; Saedler, H.; Forkmann, G.  
#authors    Eur. J. Biochem. (1993) 217:745-754  
#journal    Molecular characterization of flavanone 3-beta-hydroxylases.  
#title      Consensus sequence comparison with related enzymes and the  
              role of conserved histidine residues.  
#cross-references MUID:94039114  
#accession  S38336  
#molecule_type mRNA  
#residues     1-356 #label BR2  
#cross-references EMBL:X72593; NID:g288068; PID:g288069  
CLASSIFICATION #superfamily 1-aminocyclopropane-1-carboxylate oxidase  
KEYWORDS       oxidoreductase  
SUMMARY       #length 356 #molecular-weight 40210 #checksum 1782  
  
Query Match   31.6%; Score 6; DB 2; Length 356;  
Best Local Similarity 100.0%; Pred. No. 2.84e+00;  
Matches       6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 262 YLSNGR 267  
QY 8 YLSNGR 13  
      |||||  
RESULT 11  
ENTRY   #type complete  
TITLE   naringenin 3-dioxygenase (EC 1.14.11.9) - Bromheadia  
              finlaysoniana  
ORGANISM #formal_name Bromheadia finlaysoniana  
DATE     19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change  
12-Feb-1999  
ACCESSIONS S57750  
REFERENCE   Lim, S.H.  
#authors    Submitted to the EMBL Data Library, June 1995  
#submission  S57750  
#accession  S57750  
#molecule_type mRNA  
#residues     1-374 #label LIM  
#cross-references EMBL:X89199; NID:g895742; PID:g895743  
#experimental_source tissue type flower; clone pCF1  
GENETICS  
#gene        fht  
CLASSIFICATION #superfamily 1-aminocyclopropane-1-carboxylate oxidase  
KEYWORDS       oxidoreductase
```



```
SUMMARY          #length 374  #molecular-weight 41733  #checksum 1613

Query Match      31.6%; Score 6; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 266 YLSNGR 271
|||||
QY 8 YLSNGR 13

RESULT 12
ENTRY  JN0810  #type complete
TITLE  benzene 1,2-dioxygenase (EC 1.14.12.3) ferredoxin reductase
        component - Pseudomonas putida plasmid pHM112
ORGANISM #formal_name Pseudomonas putida
DATE 19-May-1994 #sequence_revision 06-Jan-1995 #text_change
ACCESSIONS JN0810
REFERENCE JN0810
#authors Tan, H.M.; Tang, H.Y.; Joannou, C.L.; Abdel-Wahab, N.H.;
        Mason, J.R.
#journal Gene (1993) 130:33-39
#title The Pseudomonas putida ML2 plasmid-encoded genes for benzene
        dioxygenase are unusual in codon usage and low in G+C
        content
#cross-references MUID:93345820
#accession JN0810
#molecule_type DNA
#residues 1-410 #label TAN
#experimental_source strain ML2
COMMENT This enzyme is involved in catalyzing the oxidation of benzene to
        cis-1,2-dihydroxy-cyclohexa-3,5-diene.
COMMENT This compound accepts electrons from NADH and transfers them to the
        ferredoxin component.
GENETICS
#gene bedA
#genome plasmid
CLASSIFICATION #superfamily toluene dioxygenase ferredoxin reductase
        component
KEYWORDS aromatic hydrocarbon catabolism; FAD; NAD; oxidoreductase
FEATURE
2-410 #product benzene 1,2-dioxygenase ferredoxin reductase
        component #status predicted #label MAT\
9-25 #region FAD/NAD-binding motif\
150-166 #region FAD/NAD-binding motif\
265-278 #region FAD/NAD-binding motif\
SUMMARY #length 410 #molecular-weight 43586 #checksum 3512

Query Match      31.6%; Score 6; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 359 GRIOAV 364
|||||
QY 12 GRIOAV 17

RESULT 13
ENTRY  G69123  #type complete
TITLE  conserved hypothetical protein MTH196 - Methanobacterium
        thermoautotrophicum (Strain Delta H)
ORGANISM #formal_name Methanobacterium thermoautotrophicum
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
        18-Sep-1998
ACCESSIONS G69123
REFERENCE A69000
#authors Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.;
        Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.;
        Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
        Lumm, W.; Pothier, B.; Oiu, D.; Spadafora, R.; Vicaire, R.;
        Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.; Caruso,
        A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.;

McDougall, S.; Shimer, G.; Goyal, A.; Pietrovski, S.;
Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling,
J.; Reeve, J.N.
#journal J. Bacteriol. (1997) 179:7135-7155
#title Complete genome sequence of Methanobacterium
        thermoautotrophicum Delta H: functional analysis and
        comparative genomics.
#cross-references MUID:98037514
#accession G69123
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-458 #label MTH
#cross-references GB:AE000807; GB:AE000666; NID:g2621239; PID:g2621241
#experimental_source strain Delta H
GENETICS
#gene MTH196
CLASSIFICATION #superfamily Methanobacterium thermoautotrophicum conserved
        hypothetical protein MTH196
SUMMARY #length 458 #molecular-weight 53966 #checksum 2448

Query Match      31.6%; Score 6; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 129 LSNNGRI 134
|||||
QY 9 LSNNGRI 14

RESULT 14
ENTRY  F64579  #type complete
TITLE  adenine specific DNA methyltransferase - Helicobacter pylori
        (strain 26695)
ORGANISM #formal_name Helicobacter pylori
DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change
        10-Oct-1997
ACCESSIONS F64579
REFERENCE A64520
#authors Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.;
        Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk,
        H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush,
        J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;
        Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.;
        McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.;
        Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.;
        Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.;
        Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes,
        W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
        C.M.; Venter, J.C.
#journal Nature (1997) 388:539-547
#title The complete genome sequence of the gastric pathogen
        Helicobacter pylori.
#cross-references MUID:97394467
#accession F64579
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-545 #label TOM
#cross-references GB:AE000562; GB:AE000511; NID:g2313581; PID:g2313588;
        TIGR:HP0478
SUMMARY #length 545 #molecular-weight 63177 #checksum 3160

Query Match      31.6%; Score 6; DB 2; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches          6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 511 DEKFYL 516
|||||
QY 4 DEKFYL 9

RESULT 15
```

```

ENTRY      F71932      #type complete
TITLE      type II DNA modification enzyme (methyltransferase) -
            Helicobacter pylori (strain J99)
ORGANISM   #formal_name Helicobacter pylori
            #variety   strain J99
DATE       12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
            12-Feb-1999
ACCESSIONS F71932
REFERENCE  A71800
            Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.;
            Doig, P.C.; Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge,
            B.L.; Carmel, G.; Tummino, P.J.; Caruso, A.;
            Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.;
            Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis,
            G.F.; Trust, T.J.
            Nature (1999) 397:176-180
#journal   Nature (1999) 397:176-180
#title     Genomic sequence comparison of two unrelated isolates of the
            human gastric pathogen Helicobacter pylori.
#cross-references MUID:99120557
#accession  F71932
            ##status      preliminary
            ##molecule_type DNA
            ##residues     1-545 #label ARN
            ##cross-references GB:AE001477; GB:AE001439; NID:g4154961; PID:g4154967
            ##experimental_source strain J99

GENETICS
#gene      jhp0430
SUMMARY    #length 545 #molecular-weight 63313 #checksum 5186

Query Match      31.68; Score 6; DB 2; Length 545;
Best Local Similarity 100.0%; Pred. No. 2.84e+00;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      511 DEKFYL 516
QY      4 DEKFYL 9

```

Search completed: Sat Aug 28 14:39:57 1999  
Job time : 44 secs.



97	5	26.3	408	1	BN2D_PSEPU	BENZENE 1,2-DIOXYGENAS	5.88e+01	170	5	26.3	875	1	AMD2_XENLA	PEPTIDYL-GLYCINE ALPHA	5.88e+01
98	5	26.3	409	1	TODA_PSEPU	TOLUENE 1,2-DIOXYGENAS	5.88e+01	171	5	26.3	879	1	YN65_YEAST	HYPOTHETICAL 101.7 KD	5.88e+01
99	5	26.3	417	1	YGX3_YEAST	HYPOTHETICAL 48.3 KD P	5.88e+01	172	5	26.3	882	1	YAY5_SCHPO	PROBABLE ATP-DEPENDENT	5.88e+01
100	5	26.3	423	1	ODB2_PSEPU	LIPOAMIDE ACYLTRANSFER	5.88e+01	173	5	26.3	897	1	RA16_SCHPO	DNA REPAIR PROTEIN RAD	5.88e+01
101	5	26.3	428	1	B4AR_MELGA	BETA-4C ADRENERGIC REC	5.88e+01	174	5	26.3	911	1	STEG_SCHPO	STE6 PROTEIN	5.88e+01
102	5	26.3	436	1	SYD_HALSA	ASPARTYL-TRNA SYNTHETA	5.88e+01	175	5	26.3	935	1	RNE_HAEIN	NUCLEOLAR ENZYME	5.88e+01
103	5	26.3	437	1	INTR_SACER	INTEGRASE (RECOMBINASE	5.88e+01	176	5	26.3	938	1	NM21_HUMAN	GLUTAMATE [NM2A] RECP	5.88e+01
104	5	26.3	440	1	MTN5_NOSS7	MODIFICATION 49.3 KD P	5.88e+01	177	5	26.3	938	1	NM21_MOUSE	GLUTAMATE [NM2A] RECP	5.88e+01
105	5	26.3	441	1	YXEK_BACSU	HYPOTHETICAL 49.3 KD P	5.88e+01	178	5	26.3	938	1	NM21_MOUSE	GLUTAMATE [NM2A] RECP	5.88e+01
106	5	26.3	443	1	MT17_YEAST	O-ACETYLHOMOSERINE SUL	5.88e+01	179	5	26.3	943	1	DIP2_YEAST	DOM34 INTERACTING PROT	5.88e+01
107	5	26.3	445	1	OSTA_CANFA	DOLICYL-DIPHOSPHOLIG	5.88e+01	180	5	26.3	958	1	VLA_CCMV	1A PROTEIN [CONTAINS:	5.88e+01
108	5	26.3	446	1	D3DR_RAT	D(3) DOPAMINE RECEPTOR	5.88e+01	181	5	26.3	962	1	COPE_DROME	COTONER BETA SUBUNIT	5.88e+01
109	5	26.3	446	1	D3DR_MOUSE	D(3) DOPAMINE RECEPTOR	5.88e+01	182	5	26.3	972	1	AMD_BOVIN	PEPTIDYL-GLYCINE ALPHA	5.88e+01
110	5	26.3	449	1	CPEI_MACFA	CYCLOCHROME P450 2E1 (E	5.88e+01	183	5	26.3	974	1	AMD_HUMAN	PEPTIDYL-GLYCINE ALPHA	5.88e+01
111	5	26.3	455	1	YNU6_YEAST	HYPOTHETICAL 51.6 KD P	5.88e+01	184	5	26.3	976	1	AMD_RAT	PEPTIDYL-GLYCINE ALPHA	5.88e+01
112	5	26.3	460	1	DCOR_XENLA	ORNITHINE DECARBOXYLAS	5.88e+01	185	5	26.3	979	1	AMD_MOUSE	PEPTIDYL-GLYCINE ALPHA	5.88e+01
113	5	26.3	462	1	TRPE_METHM	ANTHRANILATE SYNTHASE	5.88e+01	186	5	26.3	987	1	YD94_METJA	HYPOTHETICAL PROTEIN M	5.88e+01
114	5	26.3	464	1	TRPE_METHM	ANTHRANILATE SYNTHASE	5.88e+01	187	5	26.3	1001	1	TOP3_HUMAN	DNA TOPOISOMERASE III	5.88e+01
115	5	26.3	467	1	PSN1_HUMAN	PRESENILIN 1 (PS-1) (S	5.88e+01	188	5	26.3	1034	1	ENTK_PIG	ENTEROPEPTIDASE PRECUR	5.88e+01
116	5	26.3	482	1	K2C8_HUMAN	KERATIN, TYPE II CYTOS	5.88e+01	189	5	26.3	1041	1	DDXX_HUMAN	PUTATIVE PRE-MRNA SPLI	5.88e+01
117	5	26.3	482	1	K2C8_RAT	KERATIN, TYPE II CYTOS	5.88e+01	190	5	26.3	1050	1	ANPB_ANGJA	ATRIAL NATRIURETIC PEP	5.88e+01
118	5	26.3	484	1	GLYC_CABEL	SERINE HYDROXYMETHYLTR	5.88e+01	191	5	26.3	1102	1	RPOP_AGABT	PROBABLE DNA-DIRECTED	5.88e+01
119	5	26.3	487	1	YV14_MYCTU	HYPOTHETICAL 51.6 KD P	5.88e+01	192	5	26.3	1113	1	YV13_YEAST	HYPOTHETICAL 127.0 KD	5.88e+01
120	5	26.3	488	1	K2C8_MOUSE	KERATIN, TYPE II CYTOS	5.88e+01	193	5	26.3	1145	1	PR22_YEAST	PRE-MRNA SPLICING FACT	5.88e+01
121	5	26.3	493	1	CPEI_HUMAN	CYCLOCHROME P450 2E1 (E	5.88e+01	194	5	26.3	1251	1	YMI1_YEAST	HYPOTHETICAL 141.5 KD	5.88e+01
122	5	26.3	505	1	YPL1_YEAST	HYPOTHETICAL 58.3 KD P	5.88e+01	195	5	26.3	1372	1	FUS1_SCHPO	CELL FUSION PROTEIN FU	5.88e+01
123	5	26.3	505	1	CP57_CANNA	CYCLOCHROME P450 52D1 (E	5.88e+01	196	5	26.3	1432	1	SK13_YEAST	SUPERKILLER 3 PROTEIN.	5.88e+01
124	5	26.3	511	1	CP47_RABIT	CYCLOCHROME P450 4A7 (E	5.88e+01	197	5	26.3	1613	1	VIT2_CABEL	VITELLOGENIN 2 PRECURS	5.88e+01
125	5	26.3	516	1	P54_ENTFC	P54 PROTEIN PRECURSOR.	5.88e+01	198	5	26.3	1616	1	VIT1_CABEL	VITELLOGENIN 1 PRECURS	5.88e+01
126	5	26.3	516	1	LEGB_GOSHI	LEGUMIN B PRECURSOR (B	5.88e+01	199	5	26.3	1660	1	VITC_OSCBR	VITELLOGENIN 6 PRECURS	5.88e+01
127	5	26.3	521	1	CP51_CANNA	CYCLOCHROME P450 52A9 (	5.88e+01	200	5	26.3	1769	1	YK9_YEAST	HYPOTHETICAL 200.0 KD	5.88e+01
128	5	26.3	527	1	RF3_HAEIN	PEPTIDE CHAIN RELEASE	5.88e+01	201	5	26.3	1897	1	PTPF_HUMAN	LAR PROTEIN PRECURSOR	5.88e+01
129	5	26.3	550	1	ACAT_HUMAN	STEROL O-ACYLTRANSFER	5.88e+01	202	5	26.3	2039	1	LAR_MOUSE	PROTEIN-TYROSINE PHOSP	5.88e+01
130	5	26.3	555	1	TNSC_ECOLI	TRANSPOSIN TN7 TRANSPO	5.88e+01	203	5	26.3	2319	1	FAB_MOUSE	COAGULATION FACTOR VII	5.88e+01
131	5	26.3	556	1	ATKA_CLOAB	POTASSIUM-TRANSPORTING	5.88e+01	204	5	26.3	2325	1	PGG2_RAT	CHONDROITIN SULFATE PR	5.88e+01
132	5	26.3	559	1	QNR_COTJA	QNR-71 PROTEIN PRECURS	5.88e+01	205	5	26.3	3343	1	YOG7_CABEL	HYPOTHETICAL 375.7 KD	5.88e+01
133	5	26.3	566	1	SNFC_YEAST	TRANSCRIPTION REGULATO	5.88e+01	206	5	26.3	4447	1	PKSK_BACSU	PUTATIVE POLYKETIDE SY	5.88e+01
134	5	26.3	575	1	RM62_DROME	PUTATIVE ATP-DEPENDENT	5.88e+01	207	5	21.1	29	1	YXCX_ODOSI	HYPOTHETICAL 3.7 KD PR	1.55e+03
135	5	26.3	580	1	MTB3_BACAR	MODIFICATION METHYLASE	5.88e+01	208	4	21.1	41	1	YADH_CLOBE	HYPOTHETICAL PROTEIN I	1.55e+03
136	5	26.3	581	1	PBP2_NEIME	PENICILLIN-BINDING PRO	5.88e+01	209	4	21.1	50	1	Y476_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03
137	5	26.3	581	1	PBP2_NEIGO	PENICILLIN-BINDING PRO	5.88e+01	210	4	21.1	54	1	YJIS_ECOLI	HYPOTHETICAL 6.7 KD PR	1.55e+03
138	5	26.3	596	1	MKC7_YEAST	ASPARTIC PROTEINASE MK	5.88e+01	211	4	21.1	56	1	Y116_NPVAC	HYPOTHETICAL 6.4 KD PR	1.55e+03
139	5	26.3	602	1	ELL_MOUSE	RNA POLYMERASE II ELON	5.88e+01	212	4	21.1	59	1	TAPA_PSEFR	TEMPERATURE ACCLIMATIO	1.55e+03
140	5	26.3	606	1	KINA_BACSU	SPOREULATION KINASE A (	5.88e+01	213	4	21.1	60	1	YRHI_AZOVI	HYPOTHETICAL 7.3 KD PR	1.55e+03
141	5	26.3	608	1	UVRC_SALTY	EXCINUCLEASE ABC SUBUN	5.88e+01	214	4	21.1	64	1	YXL7_CABEL	HYPOTHETICAL 8.2 KD PR	1.55e+03
142	5	26.3	609	1	KPK1_PHAVU	PROTEIN KINASE PVPK-1	5.88e+01	215	4	21.1	67	1	Y111_NPVAC	HYPOTHETICAL 7.3 KD PR	1.55e+03
143	5	26.3	609	1	UVRC_HAEIN	EXCINUCLEASE ABC SUBUN	5.88e+01	216	4	21.1	69	1	VVDC_VACCV	HYPOTHETICAL 7.8 KD PR	1.55e+03
144	5	26.3	610	1	UVRC_ECOLI	EXCINUCLEASE ABC SUBUN	5.88e+01	217	4	21.1	71	1	YVFE_VACCC	HYPOTHETICAL 7.8 KD PR	1.55e+03
145	5	26.3	615	1	SECD_ECOLI	PROTEIN-EXPORT MEMBRAN	5.88e+01	218	4	21.1	75	1	YE05_METJA	HYPOTHETICAL 8.5 KD PR	1.55e+03
146	5	26.3	621	1	ELL_HUMAN	RNA POLYMERASE II ELON	5.88e+01	219	4	21.1	75	1	YFPU_ECOLI	HYPOTHETICAL 10.1 KD P	1.55e+03
147	5	26.3	622	1	SR68_CANFA	SIGNAL RECOGNITION PAR	5.88e+01	220	4	21.1	78	1	Y321_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03
148	5	26.3	627	1	NTG3_RAT	SODIUM- AND CHLORIDE-D	5.88e+01	221	4	21.1	79	1	Y788_METJA	HYPOTHETICAL PROTEIN M	1.55e+03
149	5	26.3	627	1	NTG3_MOUSE	SODIUM- AND CHLORIDE-D	5.88e+01	222	4	21.1	85	1	Y668_METJA	HYPOTHETICAL PROTEIN M	1.55e+03
150	5	26.3	632	1	NTG3_HUMAN	SODIUM- AND CHLORIDE-D	5.88e+01	223	4	21.1	87	1	Y060_NPVAC	HYPOTHETICAL 10.1 KD P	1.55e+03
151	5	26.3	640	1	ELL2_HUMAN	RNA POLYMERASE II ELON	5.88e+01	224	4	21.1	93	1	Y073_METJA	HYPOTHETICAL PROTEIN M	1.55e+03
152	5	26.3	643	1	HSY0_CLAHE	HEAT SHOCK 70 KD PROTE	5.88e+01	225	4	21.1	93	1	YEHE_ECOLI	HYPOTHETICAL 10.1 KD P	1.55e+03
153	5	26.3	649	1	SYM_BACST	METHIONYL-TRNA SYNTH	5.88e+01	226	4	21.1	95	1	NCCV_ALCXX	NICKEL-COBALT-CADMIUM	1.55e+03
154	5	26.3	679	1	RRP1_DROME	RECOMBINATION REPAIR P	5.88e+01	227	4	21.1	96	1	EXOX_RHIN	EXOPOLYSACCHARIDE PROD	1.55e+03
155	5	26.3	700	1	CH60_PLAEG	MITOCHONDRIAL CHAPERON	5.88e+01	228	4	21.1	98	1	EXOX_RHME	EXOPOLYSACCHARIDE PROD	1.55e+03
156	5	26.3	714	1	CLPB_MYCGE	CLPB PROTEIN.	5.88e+01	229	4	21.1	100	1	YKPB_YEAST	HYPOTHETICAL 11.4 KD P	1.55e+03
157	5	26.3	735	1	EF2_SULSO	ELONGATION FACTOR 2 (E	5.88e+01	230	4	21.1	101	1	VNS7_CVFE3	NONSTRUCTURAL PROTEIN	1.55e+03
158	5	26.3	745	1	KATE_ARATH	KINESIN-LIKE PROTEIN B	5.88e+01	231	4	21.1	101	1	VNS7_F1VP	NONSTRUCTURAL PROTEIN	1.55e+03
159	5	26.3	753	1	YJ05_CABEL	HYPOTHETICAL 85.1 KD N	5.88e+01	232	4	21.1	103	1	CYT2_MOUSE	STEFIN 2.	1.55e+03
160	5	26.3	774	1	DCMA_METJA	CARBON MONOXIDE DEHYDR	5.88e+01	233	4	21.1	105	1	YMK8_YEAST	HYPOTHETICAL 12.3 KD P	1.55e+03
161	5	26.3	778	1	TOPI_CANAL	DNA TOPOISOMERASE I (E	5.88e+01	234	4	21.1	110	1	Y12K_SMSVA	POSSIBLE 12 KD NUCLEIC	1.55e+03
162	5	26.3	796	1	PTPA_RAT	PROTEIN-TYROSINE PHOSP	5.88e+01	235	4	21.1	110	1	YV01_RABIT	IG KAPPA CHAIN V REGIO	1.55e+03
163	5	26.3	802	1	PTPA_HUMAN	PROTEIN-TYROSINE PHOSP	5.88e+01	236	4	21.1	111	1	Y101_HUMAN	HYPOTHETICAL PROTEIN K	1.55e+03
164	5	26.3	803	1	CSD2_ECOLI	OUTER MEMBRANE USHER P	5.88e+01	237	4	21.1	111	1	Y546_SYNY3	HYPOTHETICAL 11.9 KD P	1.55e+03
165	5	26.3	802	1	R1R1_CRYPV	RIBONUCLEOSIDE-DIPHOSP	5.88e+01	238	4	21.1	111	1	YGC4_YEAST	HYPOTHETICAL 12.2 KD P	1.55e+03
166	5	26.3	829	1	PTPA_MOUSE	PROTEIN-TYROSINE PHOSP	5.88e+01	239	4	21.1	115	1	YK03_CABEL	HYPOTHETICAL 13.2 KD P	1.55e+03
167	5	26.3	830	1	HMT1_SCHPO	HEAVY METAL TOLERANCE	5.88e+01	240	4	21.1	116	1	XKDD_BACSU	PHAGE-LIKE ELEMENT PBS	1.55e+03
168	5	26.3	837	1	LOX3_SOYBN	SEED LIPOXYGENASE-3 (E	5.88e+01	241	4	21.1	118	1	HV3V_HUMAN	IG HEAVY CHAIN V-III R	1.55e+03
169	5	26.3	861	1	LOX3_PEA	SEED LIPOXYGENASE-3 (E	5.88e+01	242	4	21.1	118	1	YID2_YEAST	HYPOTHETICAL 13.8 KD P	1.55e+03

243	4	21.1	1	Y341_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	316	4	21.1	1	VTIL_SCHPO	VESSICLE TRANSPORT V-SN	1.55e+03
244	4	21.1	1	YCF2_VICFA	HYPOTHETICAL ORFX PROT	1.55e+03	317	4	21.1	1	RSMB_RAT	SMALL NUCLEAR RIBONUCL	1.55e+03
245	4	21.1	1	Y008_MYCTO	VERY HYPOTHETICAL 14.0	1.55e+03	318	4	21.1	1	UBC1_YEAST	UBIQUITIN-CONJUGATING	1.55e+03
246	4	21.1	1	Y055_MYCPN	HYPOTHETICAL PROTEIN M	1.55e+03	319	4	21.1	1	HFD1_HAEIN	MINOR FIMBRIAL SUBUNIT	1.55e+03
247	4	21.1	1	YCX1_GUITH	HYPOTHETICAL 15.1 KD P	1.55e+03	320	4	21.1	1	Y443_SCHPO	HYPOTHETICAL 24.3 KD P	1.55e+03
248	4	21.1	1	YGM1_YEAST	HYPOTHETICAL 14.9 KD P	1.55e+03	321	4	21.1	1	Y4A3_BACSU	HYPOTHETICAL 25.4 KD P	1.55e+03
249	4	21.1	1	Y026_NVOP	HYPOTHETICAL 14.4 KD P	1.55e+03	322	4	21.1	1	YB8N_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03
250	4	21.1	1	H24A_CHICK	HISTONE H2A-IV	1.55e+03	323	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
251	4	21.1	1	Y266_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03	324	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
252	4	21.1	1	Y15K_SSV1	HYPOTHETICAL 14.6 KD P	1.55e+03	325	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
253	4	21.1	1	Y026_NVOP	HYPOTHETICAL 14.6 KD P	1.55e+03	326	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
254	4	21.1	1	Y0BL_BACSU	HYPOTHETICAL 15.0 KD P	1.55e+03	327	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
255	4	21.1	1	YJ8Q_ECOLI	HYPOTHETICAL 15.7 KD P	1.55e+03	328	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
256	4	21.1	1	Y454_MYCPN	HYPOTHETICAL PROTEIN M	1.55e+03	329	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
257	4	21.1	1	YKDS_BACSU	PHAGE-LIKE ELEMENT PBS	1.55e+03	330	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
258	4	21.1	1	YF32_CLOPA	HYPOTHETICAL PROTEIN I	1.55e+03	331	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
259	4	21.1	1	Y16K_TRVPS	16 KD PROTEIN	1.55e+03	332	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
260	4	21.1	1	H24X_HUMAN	HISTONE H2A X	1.55e+03	333	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
261	4	21.1	1	HBBC_ANGAN	HEMOGLOBIN CATHODIC, B	1.55e+03	334	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
262	4	21.1	1	Y220_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	335	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
263	4	21.1	1	Y180_SNNY3	HYPOTHETICAL 16.4 KD P	1.55e+03	336	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
264	4	21.1	1	Y08L_MYCTO	HYPOTHETICAL 16.8 KD P	1.55e+03	337	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
265	4	21.1	1	YAFX_ECOLI	HYPOTHETICAL 17.4 KD P	1.55e+03	338	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
266	4	21.1	1	YFJX_ECOLI	HYPOTHETICAL 17.3 KD P	1.55e+03	339	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
267	4	21.1	1	Y17K_SSV1	HYPOTHETICAL 17.8 KD P	1.55e+03	340	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
268	4	21.1	1	YR35_MYCTO	HYPOTHETICAL 16.4 KD P	1.55e+03	341	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
269	4	21.1	1	RR7_MARPO	CHLOROPLAST 30S RIBOSO	1.55e+03	342	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
270	4	21.1	1	YGBB_ECOLI	HYPOTHETICAL 16.9 KD P	1.55e+03	343	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
271	4	21.1	1	YEP7_YEAST	HYPOTHETICAL 19.0 KD P	1.55e+03	344	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
272	4	21.1	1	YRF1_SHIEL	HYPOTHETICAL 18.7 KD P	1.55e+03	345	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
273	4	21.1	1	Y400_SNNP2	HYPOTHETICAL 18.1 KD P	1.55e+03	346	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
274	4	21.1	1	YI19_MYCTO	HYPOTHETICAL 18.0 KD P	1.55e+03	347	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
275	4	21.1	1	Y005_BHP1	HYPOTHETICAL 19.2 KD P	1.55e+03	348	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
276	4	21.1	1	YF21_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03	349	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
277	4	21.1	1	YF21_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03	350	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
278	4	21.1	1	PSAN_ARATH	PHOTOSYSTEM I REACTION	1.55e+03	351	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
279	4	21.1	1	YJMA_YEAST	HYPOTHETICAL 20.3 KD P	1.55e+03	352	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
280	4	21.1	1	Y266_HAEIN	HYPOTHETICAL 20.8 KD P	1.55e+03	353	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
281	4	21.1	1	Y266_HAEIN	HYPOTHETICAL 20.8 KD P	1.55e+03	354	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
282	4	21.1	1	Y337_PORPU	HYPOTHETICAL 20.0 KD P	1.55e+03	355	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
283	4	21.1	1	Y150_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	356	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
284	4	21.1	1	YCX1_PORPU	HYPOTHETICAL 20.4 KD P	1.55e+03	357	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
285	4	21.1	1	YDXA_BACSU	HYPOTHETICAL 20.4 KD P	1.55e+03	358	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
286	4	21.1	1	YCX1_YEAST	HYPOTHETICAL 21.1 KD P	1.55e+03	359	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
287	4	21.1	1	YD21_SCHPO	HYPOTHETICAL 21.2 KD P	1.55e+03	360	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
288	4	21.1	1	RR4_SPASP	CHLOROPLAST 30S RIBOSO	1.55e+03	361	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
289	4	21.1	1	YCF4_MAIZE	HYPOTHETICAL 19.6 KD P	1.55e+03	362	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
290	4	21.1	1	YCF4_ORSEA	HYPOTHETICAL 21.5 KD P	1.55e+03	363	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
291	4	21.1	1	YNSC_MEASA	NONSTRUCTURAL PROTEIN	1.55e+03	364	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
292	4	21.1	1	YNSC_MEASA	NONSTRUCTURAL PROTEIN	1.55e+03	365	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
293	4	21.1	1	YNSC_MEASA	NONSTRUCTURAL PROTEIN	1.55e+03	366	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
294	4	21.1	1	YH20_HAEIN	HYPOTHETICAL PROTEIN H	1.55e+03	367	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
295	4	21.1	1	RR4_SIST	CHLOROPLAST 30S RIBOSO	1.55e+03	368	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
296	4	21.1	1	GREP_FRATU	GRPE PROTEIN	1.55e+03	369	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
297	4	21.1	1	MTLR_ECOLI	MANNITOL OPERON REPRES	1.55e+03	370	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
298	4	21.1	1	YAAE_BACSU	HYPOTHETICAL 21.4 KD P	1.55e+03	371	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
299	4	21.1	1	RR4_IRIPA	CHLOROPLAST 30S RIBOSO	1.55e+03	372	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
300	4	21.1	1	RR4_FESPR	CHLOROPLAST 30S RIBOSO	1.55e+03	373	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
301	4	21.1	1	RR4_ADIS	CHLOROPLAST 30S RIBOSO	1.55e+03	374	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
302	4	21.1	1	YMO7_YEAST	PUTATIVE SUCCINATE DEH	1.55e+03	375	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
303	4	21.1	1	YTAG_BACSU	HYPOTHETICAL 22.0 KD P	1.55e+03	376	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
304	4	21.1	1	HPAP_BURSO	HRP ASSOCIATED PROTEIN	1.55e+03	377	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
305	4	21.1	1	VP24_NVOP	CAPSID PROTEIN P24	1.55e+03	378	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
306	4	21.1	1	RR4_SPIOI	CHLOROPLAST 30S RIBOSO	1.55e+03	379	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
307	4	21.1	1	YMO7_YEAST	HYPOTHETICAL 23.4 KD P	1.55e+03	380	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
308	4	21.1	1	YMAA_CAEEL	HYPOTHETICAL 22.4 KD P	1.55e+03	381	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
309	4	21.1	1	Y860_RICPR	HYPOTHETICAL 23.2 KD P	1.55e+03	382	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
310	4	21.1	1	HFD2_HAEIN	MINOR FIMBRIAL SUBUNIT	1.55e+03	383	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
311	4	21.1	1	Y112_MYCGE	HYPOTHETICAL PROTEIN M	1.55e+03	384	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
312	4	21.1	1	YEO9_YEAST	HYPOTHETICAL 23.2 KD P	1.55e+03	385	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
313	4	21.1	1	YEL3_YEAST	HYPOTHETICAL 24.1 KD P	1.55e+03	386	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
314	4	21.1	1	YCD4_YEAST	HYPOTHETICAL 23.6 KD P	1.55e+03	387	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03
315	4	21.1	1	YBL5_YEAST	HYPOTHETICAL 23.9 KD P	1.55e+03	388	4	21.1	1	YDRE_SCHPO	HYPOTHETICAL 24.4 KD P	1.55e+03



535	4	21.1	409	1	YV12_MYCLE	1.55e+03	608	4	21.1	494	1	CPG1_RAT	1.55e+03
536	4	21.1	410	1	ODPX_YEAST	1.55e+03	609	4	21.1	495	1	YS02_CAEEL	1.55e+03
537	4	21.1	411	1	OMPS_VIBCH	1.55e+03	610	4	21.1	497	1	YFC7_YEAST	1.55e+03
538	4	21.1	412	1	YB83_YEAST	1.55e+03	611	4	21.1	497	1	YJ24_YEAST	1.55e+03
539	4	21.1	413	1	YBDG_ECOLI	1.55e+03	612	4	21.1	497	1	SC14_YARLI	1.55e+03
540	4	21.1	414	1	YSCD_YEREN	1.55e+03	613	4	21.1	498	1	YHJ3_ECOLI	1.55e+03
541	4	21.1	415	1	NAGA_CAEEL	1.55e+03	614	4	21.1	499	1	YH88_YEAST	1.55e+03
542	4	21.1	416	1	YKKG_CAEEL	1.55e+03	615	4	21.1	499	1	YL11_MYCHO	1.55e+03
543	4	21.1	417	1	YIBP_ECOLI	1.55e+03	616	4	21.1	499	1	KPY1_TRYBB	1.55e+03
544	4	21.1	418	1	YJ14_PYRHO	1.55e+03	617	4	21.1	500	1	YD91_SCHPO	1.55e+03
545	4	21.1	419	1	HEMA_CVBF	1.55e+03	618	4	21.1	501	1	SLAP_AEPSA	1.55e+03
546	4	21.1	420	1	YFEL_MYCMY	1.55e+03	619	4	21.1	502	1	ACH7_MOUSE	1.55e+03
547	4	21.1	421	1	YCD8_ECOLI	1.55e+03	620	4	21.1	502	1	ACH7_MOUSE	1.55e+03
548	4	21.1	422	1	OPPD_MYCPN	1.55e+03	621	4	21.1	504	1	YFGK_HAEIN	1.55e+03
549	4	21.1	423	1	HEMA_CVBM	1.55e+03	622	4	21.1	505	1	YEP4_YEAST	1.55e+03
550	4	21.1	424	1	HEMA_CVBL	1.55e+03	623	4	21.1	506	1	VP6_RDV	1.55e+03
551	4	21.1	425	1	YBYQ_YEAST	1.55e+03	624	4	21.1	507	1	YME3_YEAST	1.55e+03
552	4	21.1	426	1	YK1_CAEEL	1.55e+03	625	4	21.1	507	1	YME3_YEAST	1.55e+03
553	4	21.1	427	1	YK1_CAEEL	1.55e+03	626	4	21.1	509	1	YIFB_HAEIN	1.55e+03
554	4	21.1	428	1	Y047_METJA	1.55e+03	627	4	21.1	509	1	SYK_ACICA	1.55e+03
555	4	21.1	429	1	Y179_ARATH	1.55e+03	628	4	21.1	509	1	NCAP_SV5	1.55e+03
556	4	21.1	430	1	YQGE_BACSU	1.55e+03	629	4	21.1	511	1	YQ88_YEAST	1.55e+03
557	4	21.1	431	1	YME4_YEAST	1.55e+03	630	4	21.1	514	1	Y4UI_RHISN	1.55e+03
558	4	21.1	432	1	Y103_METJA	1.55e+03	631	4	21.1	515	1	YJEM_ECOLI	1.55e+03
559	4	21.1	433	1	YB14_YEAST	1.55e+03	632	4	21.1	515	1	YD56_SCHPO	1.55e+03
560	4	21.1	434	1	Y181_MYCPN	1.55e+03	633	4	21.1	516	1	Y067_MYCGE	1.55e+03
561	4	21.1	435	1	YSX4_CAEEL	1.55e+03	634	4	21.1	516	1	ZG53_XENLA	1.55e+03
562	4	21.1	436	1	YK88_YEAST	1.55e+03	635	4	21.1	519	1	Y771_CAEEL	1.55e+03
563	4	21.1	437	1	YCW9_YEAST	1.55e+03	636	4	21.1	519	1	VP6_WTV	1.55e+03
564	4	21.1	438	1	AP2_MOUSE	1.55e+03	637	4	21.1	519	1	VP6_WTVNJ	1.55e+03
565	4	21.1	439	1	YK21_YEAST	1.55e+03	638	4	21.1	520	1	GAL3_YEAST	1.55e+03
566	4	21.1	440	1	YJ03_ECOLI	1.55e+03	639	4	21.1	520	1	Y014_MYCTU	1.55e+03
567	4	21.1	441	1	YXEQ_BACSU	1.55e+03	640	4	21.1	523	1	YAZ4_SCHPO	1.55e+03
568	4	21.1	442	1	Y4WB_RHISN	1.55e+03	641	4	21.1	524	1	Y018_METJA	1.55e+03
569	4	21.1	443	1	Y500_PORPU	1.55e+03	642	4	21.1	526	1	YH11_YEAST	1.55e+03
570	4	21.1	444	1	HS83_LEIDO	1.55e+03	643	4	21.1	527	1	YB88_SCHPO	1.55e+03
571	4	21.1	445	1	CLG3_YEAST	1.55e+03	644	4	21.1	529	1	YH91_MYCPN	1.55e+03
572	4	21.1	446	1	YNH4_YEAST	1.55e+03	645	4	21.1	529	1	YGAA_ECOLI	1.55e+03
573	4	21.1	447	1	YK88_YEAST	1.55e+03	646	4	21.1	530	1	WC2_NEUCR	1.55e+03
574	4	21.1	448	1	Z06_XENLA	1.55e+03	647	4	21.1	530	1	YBIT_ECOLI	1.55e+03
575	4	21.1	449	1	YP95_YEAST	1.55e+03	648	4	21.1	531	1	YUTE_CAEEL	1.55e+03
576	4	21.1	450	1	YWDJ_BACSU	1.55e+03	649	4	21.1	532	1	YHCA_BACSU	1.55e+03
577	4	21.1	451	1	YAS4_HAEIN	1.55e+03	650	4	21.1	534	1	YGH_YEAST	1.55e+03
578	4	21.1	452	1	YCE8_YEAST	1.55e+03	651	4	21.1	537	1	NTPL_VARY	1.55e+03
579	4	21.1	453	1	YCE8_YEAST	1.55e+03	652	4	21.1	537	1	Y018_MYCTU	1.55e+03
580	4	21.1	454	1	YNS5_TSWV1	1.55e+03	653	4	21.1	538	1	YQ4_CAEEL	1.55e+03
581	4	21.1	455	1	YGE2_ECOLI	1.55e+03	654	4	21.1	543	1	YD83_SCHPO	1.55e+03
582	4	21.1	456	1	Y4UB_RHISN	1.55e+03	655	4	21.1	544	1	API_SCHPO	1.55e+03
583	4	21.1	457	1	YNS5_TSWV1	1.55e+03	656	4	21.1	546	1	Y772_SINY3	1.55e+03
584	4	21.1	458	1	YNS5_TSWV1	1.55e+03	657	4	21.1	546	1	Y005_SCHPO	1.55e+03
585	4	21.1	459	1	YBFW_ECOLI	1.55e+03	658	4	21.1	549	1	Y330_METJA	1.55e+03
586	4	21.1	460	1	YBFW_ECOLI	1.55e+03	659	4	21.1	550	1	HEMA_TAZH3	1.55e+03
587	4	21.1	461	1	Y064_HUMAN	1.55e+03	660	4	21.1	553	1	Y01K_ECOLI	1.55e+03
588	4	21.1	462	1	YX11_CAEEL	1.55e+03	661	4	21.1	555	1	WETA_EMENI	1.55e+03
589	4	21.1	463	1	Y215_MYCGE	1.55e+03	662	4	21.1	555	1	WETA_EMENI	1.55e+03
590	4	21.1	464	1	SG3_RAT	1.55e+03	663	4	21.1	557	1	VTFP_BPT3	1.55e+03
591	4	21.1	465	1	Z131_HUMAN	1.55e+03	664	4	21.1	558	1	Y242_METJA	1.55e+03
592	4	21.1	466	1	Y112_YEAST	1.55e+03	665	4	21.1	558	1	YNT3_YEAST	1.55e+03
593	4	21.1	467	1	HMIN_BOMMO	1.55e+03	666	4	21.1	558	1	YD2B_YEAST	1.55e+03
594	4	21.1	468	1	YNUC_DVII1	1.55e+03	667	4	21.1	560	1	YNUC_INBSI	1.55e+03
595	4	21.1	469	1	YV12_MYCTU	1.55e+03	668	4	21.1	560	1	YNUC_INBAC	1.55e+03
596	4	21.1	470	1	YAB8_SCHPO	1.55e+03	669	4	21.1	560	1	YNUC_INBAC	1.55e+03
597	4	21.1	471	1	Y740_HAEIN	1.55e+03	670	4	21.1	560	1	YNUC_INBAC	1.55e+03
598	4	21.1	472	1	ST13_SCHPO	1.55e+03	671	4	21.1	560	1	YNUC_INBAA	1.55e+03
599	4	21.1	473	1	YB10_METJA	1.55e+03	672	4	21.1	560	1	YEGC_YEAST	1.55e+03
600	4	21.1	474	1	YU42_MYCTU	1.55e+03	673	4	21.1	560	1	YU42_MYCTU	1.55e+03
601	4	21.1	475	1	Y118_YEAST	1.55e+03	674	4	21.1	563	1	YU42_YEAST	1.55e+03
602	4	21.1	476	1	YU42_YEAST	1.55e+03	675	4	21.1	565	1	YU42_YEAST	1.55e+03
603	4	21.1	477	1	YU42_YEAST	1.55e+03	676	4	21.1	566	1	YU42_YEAST	1.55e+03
604	4	21.1	478	1	YU42_YEAST	1.55e+03	677	4	21.1	566	1	YU42_YEAST	1.55e+03
605	4	21.1	479	1	YU42_YEAST	1.55e+03	678	4	21.1	566	1	YU42_YEAST	1.55e+03
606	4	21.1	480	1	YU42_YEAST	1.55e+03	679	4	21.1	568	1	YU42_YEAST	1.55e+03
607	4	21.1	481	1	YU42_YEAST	1.55e+03	680	4	21.1	569	1	YU42_YEAST	1.55e+03

681	4	21.1	569	1	Y139_MYCGE	HYPOTHETICAL PROTEIN M	1.55e+03	754	4	21.1	716	1	RRP2_TAZTF	RNA-DIRECTED RNA POLYM	1.55e+03
682	4	21.1	569	1	UREL_SNY3	UREASE ALPHA SUBUNIT (	1.55e+03	755	4	21.1	718	1	FLGE_HELPT	FLAGELLAR HOOK PROTEIN	1.55e+03
683	4	21.1	571	1	YHV4_YEAST	HYPOTHETICAL 65.8 KD P	1.55e+03	756	4	21.1	721	1	YCF2_OENPI	HYPOTHETICAL PROTEIN F	1.55e+03
684	4	21.1	571	1	UREL_YERPS	UREASE ALPHA SUBUNIT (	1.55e+03	757	4	21.1	725	1	MYT1_HUMAN	MYELIN TRANSCRIPTION F	1.55e+03
685	4	21.1	574	1	YFHB_BORPE	HYPOTHETICAL PROTEIN I	1.55e+03	758	4	21.1	725	1	YGL4_YEAST	HYPOTHETICAL 80.0 KD P	1.55e+03
686	4	21.1	574	1	YF3_CAEEL	PUTATIVE SERINE CARBOX	1.55e+03	759	4	21.1	726	1	Z184_HUMAN	ZINC FINGER PROTEIN 18	1.55e+03
687	4	21.1	574	1	SFCA_ECOLI	PROBABLE MALATE OXIDOR	1.55e+03	760	4	21.1	728	1	YJ89_YEAST	HYPOTHETICAL 85.0 KD P	1.55e+03
688	4	21.1	574	1	YJ23_YEAST	HYPOTHETICAL 66.1 KD P	1.55e+03	761	4	21.1	732	1	VP4_WTV	NONSTRUCTURAL PROTEIN	1.55e+03
689	4	21.1	579	1	Y07D_MYCTU	HYPOTHETICAL ABC TRANS	1.55e+03	762	4	21.1	735	1	YKCC_CAEEL	HYPOTHETICAL 83.7 KD P	1.55e+03
690	4	21.1	579	1	YKMA_YEAST	HYPOTHETICAL 65.1 KD P	1.55e+03	763	4	21.1	740	1	YCF2_CUSRE	HYPOTHETICAL 85.4 KD P	1.55e+03
691	4	21.1	579	1	G160_HUMAN	GOLGIN-160 (FRAGMENT)	1.55e+03	764	4	21.1	740	1	YXGF_ECOLI	HYPOTHETICAL 81.4 KD P	1.55e+03
692	4	21.1	583	1	YCV1_YEAST	HYPOTHETICAL 63.0 KD P	1.55e+03	765	4	21.1	741	1	LOXB_PHAVU	LIPOXYGENASE (EC 1.13.	1.55e+03
693	4	21.1	586	1	Y4GM_RHISN	PROBABLE ABC TRANSPORT	1.55e+03	766	4	21.1	741	1	IDH_AZOVI	ISOCITRATE DEHYDROGENA	1.55e+03
694	4	21.1	587	1	YKK0_YEAST	HYPOTHETICAL 67.5 KD P	1.55e+03	767	4	21.1	743	1	YK47_YEAST	HYPOTHETICAL 83.6 KD P	1.55e+03
695	4	21.1	589	1	Y015_MYCGE	HYPOTHETICAL ABC TRANS	1.55e+03	768	4	21.1	744	1	YGV4_HALSO	HYPOTHETICAL 80.2 KD P	1.55e+03
696	4	21.1	590	1	YNN7_YEAST	HYPOTHETICAL 68.8 KD P	1.55e+03	769	4	21.1	748	1	Y875_METJA	HYPOTHETICAL PROTEIN M	1.55e+03
697	4	21.1	594	1	Y702_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	770	4	21.1	749	1	JSB2_TRINI	BASIC JUVENILE HORMONE	1.55e+03
698	4	21.1	594	1	YRN8_CAEEL	HYPOTHETICAL 66.9 KD P	1.55e+03	771	4	21.1	751	1	YK09_YEAST	HYPOTHETICAL 85.5 KD P	1.55e+03
699	4	21.1	595	1	YCB6_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	772	4	21.1	751	1	YLN5_CAEEL	HYPOTHETICAL 84.8 KD P	1.55e+03
700	4	21.1	596	1	TR4_MOUSE	ORPHAN RECEPTOR TR4 (O	1.55e+03	773	4	21.1	752	1	KS6A_CHICK	REBOSOMAL PROTEIN S6 K	1.55e+03
701	4	21.1	596	1	TR4_HUMAN	ORPHAN RECEPTOR TR4 (O	1.55e+03	774	4	21.1	753	1	NEC1_HUMAN	NEUROENDOCRINE CONVERT	1.55e+03
702	4	21.1	597	1	YK4_KLUJA	HYPOTHETICAL KILLER PL	1.55e+03	775	4	21.1	754	1	ECEL_BOVIN	ENDOTHELIN-CONVERTING	1.55e+03
703	4	21.1	599	1	YJ01_YEAST	HYPOTHETICAL 68.5 KD P	1.55e+03	776	4	21.1	756	1	Y328_MYCGE	HYPOTHETICAL PROTEIN M	1.55e+03
704	4	21.1	603	1	PLK1_MOUSE	SERINE/THREONINE-PROTE	1.55e+03	777	4	21.1	759	1	YEHM_ECOLI	HYPOTHETICAL 83.4 KD P	1.55e+03
705	4	21.1	604	1	YJY5_CAEEL	PROBABLE G PROTEIN-COU	1.55e+03	778	4	21.1	766	1	Y713_METJA	PUTATIVE HYDROGENASE E	1.55e+03
706	4	21.1	605	1	YG2S_YEAST	HYPOTHETICAL 69.9 KD P	1.55e+03	779	4	21.1	767	1	YNY7_YEAST	PUTATIVE CYSTEINYL-TRN	1.55e+03
707	4	21.1	607	1	YAP7_SCHPO	HYPOTHETICAL 68.8 KD P	1.55e+03	780	4	21.1	767	1	TOP1_CRIGR	DNA TOPOISOMERASE I (E	1.55e+03
708	4	21.1	607	1	VPR_BHP1	PROBABLE TERMINASE, AT	1.55e+03	781	4	21.1	768	1	TRPG_ASPAW	ANTHRANILATE SYNTHASE	1.55e+03
709	4	21.1	607	1	YN92_YEAST	PUTATIVE TRANSCRIPTION	1.55e+03	782	4	21.1	769	1	SWI6_KLUJA	REGULATORY PROTEIN SWI	1.55e+03
710	4	21.1	608	1	YD93_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	783	4	21.1	776	1	VP4_ROTTH	OUTER CAPSID PROTEIN V	1.55e+03
711	4	21.1	616	1	NDVA_RHIME	BETA-(1-->2)GLUCAN EXP	1.55e+03	784	4	21.1	776	1	VP4_ROTFF	OUTER CAPSID PROTEIN V	1.55e+03
712	4	21.1	616	1	SPAN_STRPU	SPAN PROTEIN PRECURSOR	1.55e+03	785	4	21.1	776	1	VP4_ROT9	OUTER CAPSID PROTEIN V	1.55e+03
713	4	21.1	617	1	YACH_ECOLI	HYPOTHETICAL 69.4 KD P	1.55e+03	786	4	21.1	776	1	VP4_ROTTH	OUTER CAPSID PROTEIN V	1.55e+03
714	4	21.1	618	1	YK3_YEAST	PUTATIVE 70.4 KD TRANS	1.55e+03	787	4	21.1	776	1	VP4_ROTTP	OUTER CAPSID PROTEIN V	1.55e+03
715	4	21.1	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	1.55e+03	788	4	21.1	776	1	VP4_ROTTP	OUTER CAPSID PROTEIN V	1.55e+03
716	4	21.1	622	1	YFEL_YEAST	PUTATIVE MULTICOPPER O	1.55e+03	789	4	21.1	776	1	YK5_CAEEL	HYPOTHETICAL 89.0 KD P	1.55e+03
717	4	21.1	623	1	Y014_MYCPN	HYPOTHETICAL ABC TRANS	1.55e+03	790	4	21.1	776	1	KLPI_CHLRE	KINESIN-LIKE PROTEIN K	1.55e+03
718	4	21.1	623	1	Y014_MYCGE	HYPOTHETICAL ABC TRANS	1.55e+03	791	4	21.1	778	1	YF05_METJA	PUTATIVE ATP-DEPENDENT	1.55e+03
719	4	21.1	629	1	YD4B_SCHPO	HYPOTHETICAL 71.3 KD P	1.55e+03	792	4	21.1	779	1	K6PF_HUMAN	6-PHOSPHOFRUCTOKINASE	1.55e+03
720	4	21.1	630	1	YCF2_OENVL	HYPOTHETICAL PROTEIN (	1.55e+03	793	4	21.1	783	1	YAF3_YEAST	HYPOTHETICAL 87.5 KD P	1.55e+03
721	4	21.1	630	1	Y4B3_RHISN	HYPOTHETICAL 67.9 KD P	1.55e+03	794	4	21.1	783	1	ZFY2_MOUSE	ZINC FINGER Y-CHROMOSO	1.55e+03
722	4	21.1	631	1	YC10_ECOLI	HYPOTHETICAL 71.3 KD P	1.55e+03	795	4	21.1	783	1	GARN_DROME	GARNET PROTEIN.	1.55e+03
723	4	21.1	631	1	NTPI_VACCV	NUCLEOSIDE TRIPHOSPHAT	1.55e+03	796	4	21.1	785	1	YE15_CAEEL	HYPOTHETICAL 90.8 KD P	1.55e+03
724	4	21.1	632	1	YK65_CAEEL	HYPOTHETICAL 71.9 KD P	1.55e+03	797	4	21.1	785	1	Y043_METJA	HYPOTHETICAL PROTEIN M	1.55e+03
725	4	21.1	632	1	Y041_HUMAN	HYPOTHETICAL PROTEIN K	1.55e+03	798	4	21.1	786	1	Y09S_MYCTU	HYPOTHETICAL 87.3 KD P	1.55e+03
726	4	21.1	633	1	YH84_YEAST	HYPOTHETICAL 71.2 KD P	1.55e+03	799	4	21.1	789	1	Y38C_MYCTU	HYPOTHETICAL 69.2 KD A	1.55e+03
727	4	21.1	637	1	MUTA_PROFR	METHYLMALONYL-COA MUTA	1.55e+03	800	4	21.1	790	1	YDDB_ECOLI	HYPOTHETICAL 89.3 KD P	1.55e+03
728	4	21.1	638	1	NEC2_PIG	NEUROENDOCRINE CONVERT	1.55e+03	801	4	21.1	792	1	Y09S_MYCLE	HYPOTHETICAL 88.1 KD P	1.55e+03
729	4	21.1	638	1	DNAK_RHILE	DNAP PROTEIN (HEAT SHO	1.55e+03	802	4	21.1	793	1	STA5_RAT	SIGNAL TRANSDUCER AND	1.55e+03
730	4	21.1	640	1	YIF6_YEAST	HYPOTHETICAL 71.4 KD P	1.55e+03	803	4	21.1	796	1	YH04_YEAST	HYPOTHETICAL 91.2 KD P	1.55e+03
731	4	21.1	647	1	MAOC_FLAPR	MALATE OXIDOREDUCTASE,	1.55e+03	804	4	21.1	796	1	YAB9_SCHPO	HYPOTHETICAL 90.9 KD P	1.55e+03
732	4	21.1	655	1	YKDA_MYCCA	HYPOTHETICAL 75.9 KD P	1.55e+03	805	4	21.1	797	1	YB92_YEAST	HYPOTHETICAL 92.9 KD P	1.55e+03
733	4	21.1	656	1	YC36_PORPU	HYPOTHETICAL SENSOR-LI	1.55e+03	806	4	21.1	798	1	PHSG_BACSU	GLYCOGEN PHOSPHORYLASE	1.55e+03
734	4	21.1	657	1	YG66_YEAST	HYPOTHETICAL 72.9 KD P	1.55e+03	807	4	21.1	803	1	ZN43_HUMAN	ZINC FINGER PROTEIN 43	1.55e+03
735	4	21.1	663	1	YJ09_YEAST	HYPOTHETICAL 77.4 KD P	1.55e+03	808	4	21.1	807	1	AFSK_STRGR	SERINE/THREONINE PROTE	1.55e+03
736	4	21.1	671	1	RECG_STRPN	ATP-DEPENDENT DNA HELI	1.55e+03	809	4	21.1	815	1	PHSG_ECOLI	GLYCOGEN PHOSPHORYLASE	1.55e+03
737	4	21.1	672	1	Y32B_MYCPN	HYPOTHETICAL PROTEIN M	1.55e+03	810	4	21.1	817	1	RRPO_CNV	PROBABLE RNA-DIRECTED	1.55e+03
738	4	21.1	672	1	VNCS_PAVHH	NONCAPSID PROTEIN NS-1	1.55e+03	811	4	21.1	824	1	YOT5_CAEEL	HYPOTHETICAL 95.7 KD P	1.55e+03
739	4	21.1	672	1	VNCS_MOMIM	NONCAPSID PROTEIN NS-1	1.55e+03	812	4	21.1	828	1	YFA4_YEAST	HYPOTHETICAL 95.4 KD P	1.55e+03
740	4	21.1	676	1	YMEH_CAEEL	YME1 PROTEIN HOMOLOG (	1.55e+03	813	4	21.1	832	1	YFC4_YEAST	HYPOTHETICAL 96.7 KD P	1.55e+03
741	4	21.1	676	1	HS70_BREL	HEAT SHOCK 70 KD PROTE	1.55e+03	814	4	21.1	832	1	YC26_YEAST	PUTATIVE 95.7 KD TRANS	1.55e+03
742	4	21.1	679	1	YHC9_YEAST	HYPOTHETICAL 77.8 KD P	1.55e+03	815	4	21.1	841	1	YAGX_ECOLI	HYPOTHETICAL 91.2 KD P	1.55e+03
743	4	21.1	681	1	MAOC_ECOLI	MAOC PROTEIN (PHENYLAC	1.55e+03	816	4	21.1	842	1	YJE7_YEAST	HYPOTHETICAL 99.3 KD P	1.55e+03
744	4	21.1	687	1	YBQ0_YEAST	HYPOTHETICAL 78.8 KD P	1.55e+03	817	4	21.1	847	1	PHS1_HUMAN	GLYCOGEN PHOSPHORYLASE	1.55e+03
745	4	21.1	691	1	YHUG_ECOLI	HYPOTHETICAL 75.1 KD P	1.55e+03	818	4	21.1	848	1	YBVB_YEAST	HYPOTHETICAL 92.8 KD P	1.55e+03
746	4	21.1	692	1	YI13_CAEEL	PROBABLE G PROTEIN-COU	1.55e+03	819	4	21.1	852	1	YHQ3_YEAST	HYPOTHETICAL 96.4 KD P	1.55e+03
747	4	21.1	697	1	Y0BY_MYCTU	HYPOTHETICAL PROTEIN C	1.55e+03	820	4	21.1	854	1	VP41_HUMAN	VACUOLAR ASSEMBLY PROT	1.55e+03
748	4	21.1	703	1	YG52_HUMAN	GLYCOGEN [STARCH] SYNT	1.55e+03	821	4	21.1	855	1	HAL5_YEAST	PROTEIN KINASE HAL5 (E	1.55e+03
749	4	21.1	704	1	YG29_YEAST	HYPOTHETICAL 82.8 KD P	1.55e+03	822	4	21.1	858	1	YNU1_YEAST	HYPOTHETICAL 98.1 KD P	1.55e+03
750	4	21.1	713	1	YK2_MOUSE	ZONA PELLUCIDA SPERM-B	1.55e+03	823	4	21.1	861	1	NASA_KLEEN	NITRATE REDUCTASE (EC	1.55e+03
751	4	21.1	713	1	YH04_YEAST	HYPOTHETICAL 80.1 KD P	1.55e+03	824	4	21.1	862	1	LOX1_HORVU	LIPOXYGENASE 1 (EC 1.1	1.55e+03
752	4	21.1	714	1	YFC4_ECOLI	PUTATIVE FATTY OXIDATI	1.55e+03	825	4	21.1	863	1	YEJO_ECOLI	HYPOTHETICAL 91.2 KD P	1.55e+03
753	4	21.1	715	1	ERF2_CANAL	EUKARYOTIC PEPTIDE CHA	1.55e+03	826	4	21.1	865	1	VU5_HSV7J	U5 PROTEIN.	1.55e+03



827	4	21.1	866	1	YCBS_ECOLI	HYPOTHEICAL OUTER MEM	1.55e+03	900	4	21.1	1070	1	YHVA_YEAST	HYPOTHEICAL 123.0 KD	1.55e+03
828	4	21.1	872	1	VP2_RORPC	RNA-BINDING PROTEIN VP	1.55e+03	901	4	21.1	1070	1	AGLU_CANTS	ALPHA-GLUCOSIDASE PREC	1.55e+03
829	4	21.1	873	1	YJ34_YEAST	HYPOTHEICAL 97.1 KD P	1.55e+03	902	4	21.1	1073	1	YASS_SCHPO	HYPOTHEICAL 120.6 KD	1.55e+03
830	4	21.1	874	1	SLAP_BACLI	S-LAYER PROTEIN PRECUR	1.55e+03	903	4	21.1	1079	1	IF2P_SCHPO	PROBABLE TRANSLATION I	1.55e+03
831	4	21.1	881	1	YJH8_YEAST	HYPOTHEICAL 89.2 KD P	1.55e+03	904	4	21.1	1080	1	SET1_YEAST	SET1 PROTEIN	1.55e+03
832	4	21.1	881	1	YH72_HAEN	HYPOTHEICAL PROTEIN H	1.55e+03	905	4	21.1	1083	1	YIL2_YEAST	HYPOTHEICAL 123.6 KD	1.55e+03
833	4	21.1	883	1	HMDH_XENLA	3-HYDROXY-3-METHYLGUT	1.55e+03	906	4	21.1	1102	1	YK54_CAEEL	HYPOTHEICAL 124.8 KD	1.55e+03
834	4	21.1	885	1	Y143_HUMAN	HYPOTHEICAL PROTEIN K	1.55e+03	907	4	21.1	1107	1	YLR2_CAEEL	HYPOTHEICAL 122.7 KD	1.55e+03
835	4	21.1	887	1	Y277_MYCPN	HYPOTHEICAL PROTEIN M	1.55e+03	908	4	21.1	1114	1	YAX6_SCHPO	HYPOTHEICAL 126.5 KD	1.55e+03
836	4	21.1	888	1	HPDH_RABIT	3-HYDROXY-3-METHYLGUT	1.55e+03	909	4	21.1	1114	1	YR18_YEAST	DNA REPAIR PROTEIN RHC	1.55e+03
837	4	21.1	889	1	RMDL_METVA	DNA-DIRECTED RNA POLYM	1.55e+03	910	4	21.1	1118	1	YPT4_CAEEL	HYPOTHEICAL 127.3 KD	1.55e+03
838	4	21.1	892	1	YJ06_YEAST	HYPOTHEICAL 103.5 KD	1.55e+03	911	4	21.1	1118	1	YIP1_YEAST	HYPOTHEICAL 128.7 KD	1.55e+03
839	4	21.1	895	1	YHED_ECOLI	HYPOTHEICAL PROTEIN M	1.55e+03	912	4	21.1	1120	1	YB85_YEAST	HYPOTHEICAL 124.0 KD	1.55e+03
840	4	21.1	897	1	XPC_MOUSE	DNA-REPAIR PROTEIN COM	1.55e+03	913	4	21.1	1121	1	YJ09_YEAST	HYPOTHEICAL 127.4 KD	1.55e+03
841	4	21.1	900	1	VP3_BT1V3	VP3 CORE PROTEIN (MAJO	1.55e+03	914	4	21.1	1121	1	YEO6_YEAST	HYPOTHEICAL 127.0 KD	1.55e+03
842	4	21.1	901	1	VP3_BT1V1	VP3 CORE PROTEIN (MAJO	1.55e+03	915	4	21.1	1125	1	Y079_HUMAN	HYPOTHEICAL PROTEIN K	1.55e+03
843	4	21.1	901	1	VP3_BT1V1	VP3 CORE PROTEIN (MAJO	1.55e+03	916	4	21.1	1129	1	YB95_YEAST	HYPOTHEICAL 131.1 KD	1.55e+03
844	4	21.1	901	1	VP3_BT1V1	VP3 CORE PROTEIN (MAJO	1.55e+03	917	4	21.1	1139	1	Y060_CAEEL	HYPOTHEICAL 128.6 KD	1.55e+03
845	4	21.1	901	1	VP3_BT1V1	VP3 CORE PROTEIN (MAJO	1.55e+03	918	4	21.1	1139	1	VRNA_BSMV	ALPHA-A PROTEIN	1.55e+03
846	4	21.1	901	1	VP3_BT1V7	VP3 CORE PROTEIN (MAJO	1.55e+03	919	4	21.1	1140	1	YHOL_ECOLI	HYPOTHEICAL 128.7 KD	1.55e+03
847	4	21.1	901	1	VP3_BT1V0	VP3 CORE PROTEIN (MAJO	1.55e+03	920	4	21.1	1174	1	YJ11_YEAST	HYPOTHEICAL 135.1 KD	1.55e+03
848	4	21.1	901	1	VP3_BT1V2A	VP3 CORE PROTEIN (MAJO	1.55e+03	921	4	21.1	1175	1	SRG2_YEAST	ATP-DEPENDENT DNA HELI	1.55e+03
849	4	21.1	904	1	Y4G1_RHSN	HYPOTHEICAL 102.8 KD	1.55e+03	922	4	21.1	1183	1	YK6_YEAST	HYPOTHEICAL 133.3 KD	1.55e+03
850	4	21.1	909	1	Y4G1_RHSN	HYPOTHEICAL 102.8 KD	1.55e+03	923	4	21.1	1197	1	Y4CA_RHSN	HYPOTHEICAL 133.7 KD	1.55e+03
851	4	21.1	912	1	UBP3_YEAST	UBIQUITIN CARBOXYL-TER	1.55e+03	924	4	21.1	1206	1	Y054_MYCGE	HYPOTHEICAL PROTEIN M	1.55e+03
852	4	21.1	916	1	TOPI_ARATH	DNA TOPOISOMERASE I (E	1.55e+03	925	4	21.1	1219	1	YM29_YEAST	HYPOTHEICAL MYOSIN-LI	1.55e+03
853	4	21.1	921	1	ITH4_PIG	INTER-ALPHA-TRYPSIN IN	1.55e+03	926	4	21.1	1228	1	YH05_YEAST	HYPOTHEICAL 143.6 KD	1.55e+03
854	4	21.1	922	1	YB1C_SCHPO	HYPOTHEICAL 103.4 KD	1.55e+03	927	4	21.1	1232	1	YQV5_YEAST	HYPOTHEICAL 134.9 KD	1.55e+03
855	4	21.1	923	1	YOL3_CAEEL	HYPOTHEICAL 105.8 KD	1.55e+03	928	4	21.1	1233	1	YF16_YEAST	HYPOTHEICAL 137.7 KD	1.55e+03
856	4	21.1	928	1	YLR3_CAEEL	HYPOTHEICAL 106.7 KD	1.55e+03	929	4	21.1	1233	1	VL1_RGOVD	MAJOR CORE PROTEIN LAM	1.55e+03
857	4	21.1	928	1	YG23_YEAST	HYPOTHEICAL 106.7 KD	1.55e+03	930	4	21.1	1238	1	YQ09_CAEEL	HYPOTHEICAL 141.2 KD	1.55e+03
858	4	21.1	939	1	XPC_HUMAN	DNA-REPAIR PROTEIN COM	1.55e+03	931	4	21.1	1250	1	YFAL_ECOLI	HYPOTHEICAL 143.3 KD	1.55e+03
859	4	21.1	941	1	YR01_CAMJE	HYPOTHEICAL 107.7 KD	1.55e+03	932	4	21.1	1253	1	YAGD_SCHPO	HYPOTHEICAL 143.3 KD	1.55e+03
860	4	21.1	950	1	MIC1_YEAST	MIC1 PROTEIN	1.55e+03	933	4	21.1	1254	1	YKCA_CAEEL	HYPOTHEICAL 142.5 KD	1.55e+03
861	4	21.1	955	1	VP2_BT1V7	OUTER CAPSID PROTEIN V	1.55e+03	934	4	21.1	1273	1	YAR2_SCHPO	HYPOTHEICAL 142.5 KD	1.55e+03
862	4	21.1	956	1	GLK4_RAT	OUTER CAPSID PROTEIN V	1.55e+03	935	4	21.1	1275	1	YAU9_SCHPO	HYPOTHEICAL 143.6 KD	1.55e+03
863	4	21.1	956	1	YKRI_CAEEL	GLUTAMATE RECEPTOR, IO	1.55e+03	936	4	21.1	1290	1	VACO_HELPY	VACUOLAIN CYTOTOXIN	1.55e+03
864	4	21.1	957	1	YKRI_CAEEL	HYPOTHEICAL 112.1 KD	1.55e+03	937	4	21.1	1318	1	YPI4_EBV	PROBABLE MEMBRANE ANTI	1.55e+03
865	4	21.1	958	1	YKRI_CAEEL	HYPOTHEICAL 108.2 KD	1.55e+03	938	4	21.1	1325	1	YDEK_ECOLI	XANTHINE DEHYDROGENASE	1.55e+03
866	4	21.1	959	1	VP2_BT1V3V	OUTER CAPSID PROTEIN V	1.55e+03	939	4	21.1	1330	1	YKRI_CAEEL	XANTHINE DEHYDROGENASE	1.55e+03
867	4	21.1	960	1	VP41_LYCES	VACUOLAR ASSEMBLY PROT	1.55e+03	940	4	21.1	1331	1	YKRI_CAEEL	XANTHINE DEHYDROGENASE	1.55e+03
868	4	21.1	960	1	VP41_LYCES	VACUOLAR ASSEMBLY PROT	1.55e+03	941	4	21.1	1332	1	YKRI_CAEEL	XANTHINE DEHYDROGENASE	1.55e+03
869	4	21.1	964	1	DPOL_CBEPV	DNA POLYMERASE (EC 2.7	1.55e+03	942	4	21.1	1335	1	YKRI_CAEEL	XANTHINE DEHYDROGENASE	1.55e+03
870	4	21.1	969	1	Y288_HUMAN	HYPOTHEICAL PROTEIN K	1.55e+03	943	4	21.1	1337	1	YDM5_SCHPO	HYPOTHEICAL 150.8 KD	1.55e+03
871	4	21.1	969	1	Y288_HUMAN	HYPOTHEICAL PROTEIN K	1.55e+03	944	4	21.1	1345	1	YH00_YEAST	HYPOTHEICAL 149.7 KD	1.55e+03
872	4	21.1	970	1	Y277_MYCGE	PUTATIVE ZINC-PROTEASE	1.55e+03	945	4	21.1	1358	1	YKRI_CAEEL	XANTHINE DEHYDROGENASE	1.55e+03
873	4	21.1	970	1	NAC1_MOUSE	HYPOTHEICAL PROTEIN M	1.55e+03	946	4	21.1	1361	1	YME9_YEAST	XANTHINE DEHYDROGENASE	1.55e+03
874	4	21.1	971	1	NAC1_MOUSE	SODIUM/CALCIUM EXCHANG	1.55e+03	947	4	21.1	1380	1	ZMS1_YEAST	ZINC FINGER PROTEIN 2M	1.55e+03
875	4	21.1	980	1	GLK5_HUMAN	GLUTAMATE RECEPTOR, IO	1.55e+03	948	4	21.1	1385	1	YMS5_CAEEL	HYPOTHEICAL 159.2 KD	1.55e+03
876	4	21.1	981	1	YKRI_CAEEL	HYPOTHEICAL 113.2 KD	1.55e+03	949	4	21.1	1391	1	YKRI_CAEEL	DNA-DIRECTED RNA POLYM	1.55e+03
877	4	21.1	987	1	EPB4_MOUSE	EPHRIIN TYPE-B RECEPTOR	1.55e+03	950	4	21.1	1401	1	WRN_MOUSE	WERNER SYNDROME HELICA	1.55e+03
878	4	21.1	987	1	EPB4_MOUSE	EPHRIIN TYPE-B RECEPTOR	1.55e+03	951	4	21.1	1403	1	YD53_SCHPO	PROBABLE EUKARYOTIC IN	1.55e+03
879	4	21.1	988	1	YKRI_CAEEL	HYPOTHEICAL 112.2 KD	1.55e+03	952	4	21.1	1411	1	YKRI_CAEEL	HYPOTHEICAL 162.7 KD	1.55e+03
880	4	21.1	989	1	YKRI_CAEEL	HYPOTHEICAL 115.3 KD	1.55e+03	953	4	21.1	1416	1	YKRI_CAEEL	HYPOTHEICAL 316.1 KD	1.55e+03
881	4	21.1	992	1	MA2B_MOUSE	LYSOSOMAL ALPHA-MANNOS	1.55e+03	954	4	21.1	1459	1	YF1M_CAEEL	HYPOTHEICAL 166.0 KD	1.55e+03
882	4	21.1	993	1	YD43_SCHPO	HYPOTHEICAL 111.9 KD	1.55e+03	955	4	21.1	1489	1	YKRI_CAEEL	HYPOTHEICAL 171.5 KD	1.55e+03
883	4	21.1	993	1	NISB_LACLA	NISIN BIOSYNTHESIS PRO	1.55e+03	956	4	21.1	1499	1	CDR2_CANAL	MULTIDRUG RESISTANCE P	1.55e+03
884	4	21.1	993	1	EPB2_MOUSE	EPHRIIN TYPE-B RECEPTOR	1.55e+03	957	4	21.1	1569	1	YKRI_CAEEL	HYPOTHEICAL 98.4 KD P	1.55e+03
885	4	21.1	995	1	YKRI_CAEEL	HYPOTHEICAL 111.2 KD	1.55e+03	958	4	21.1	1583	1	YKRI_CAEEL	HYPOTHEICAL 180.2 KD	1.55e+03
886	4	21.1	997	1	YKRI_CAEEL	HYPOTHEICAL 110.3 KD	1.55e+03	959	4	21.1	1592	1	YH05_YEAST	PROBABLE ATP-DEPENDENT	1.55e+03
887	4	21.1	1005	1	YKRI_CAEEL	HYPOTHEICAL 110.3 KD	1.55e+03	960	4	21.1	1592	1	YKRI_CAEEL	HYPOTHEICAL 178.4 KD	1.55e+03
888	4	21.1	1010	1	MA2B_HUMAN	P53-BINDING PROTEIN 53	1.55e+03	961	4	21.1	1616	1	YKRI_CAEEL	HYPOTHEICAL PROTEIN C	1.55e+03
889	4	21.1	1018	1	YK26_CAEEL	LYSOSOMAL ALPHA-MANNOS	1.55e+03	962	4	21.1	1658	1	YKRI_CAEEL	HYPOTHEICAL 187.1 KD	1.55e+03
890	4	21.1	1019	1	YKRI_CAEEL	MAJOR 114 KD STRUCTURA	1.55e+03	963	4	21.1	1666	1	YKRI_CAEEL	PROBABLE CLATHRIN HEAV	1.55e+03
891	4	21.1	1020	1	YKRI_CAEEL	HYPOTHEICAL 112.3 KD	1.55e+03	964	4	21.1	1679	1	YKRI_CAEEL	HYPOTHEICAL 195.1 KD	1.55e+03
892	4	21.1	1021	1	YKRI_CAEEL	HYPOTHEICAL 111.9 KD	1.55e+03	965	4	21.1	1687	1	YKRI_CAEEL	ZINC FINGER PROTEIN 14	1.55e+03
893	4	21.1	1033	1	YKRI_CAEEL	HYPOTHEICAL PROTEIN M	1.55e+03	966	4	21.1	1694	1	YKRI_CAEEL	ZINC FINGER HEAVY CHAIN	1.55e+03
894	4	21.1	1035	1	POL_HV28B	POLYPROTEIN [CONTA	1.55e+03	967	4	21.1	1722	1	YKRI_CAEEL	HYPOTHEICAL HELICASE	1.55e+03
895	4	21.1	1044	1	YKRI_CAEEL	HYPOTHEICAL PROTEIN C	1.55e+03	968	4	21.1	1744	1	TENS_CHICK	TENSIN	1.55e+03
896	4	21.1	1056	1	SPS_SPTOL	SUCROSE-PHOSPHATE SYNT	1.55e+03	969	4	21.1	1748	1	YKRI_CAEEL	HYPOTHEICAL 196.1 KD	1.55e+03
897	4	21.1	1057	1	YKRI_CAEEL	DORSAL-VENTRAL PATTERN	1.55e+03	970	4	21.1	1750	1	YKRI_CAEEL	HYPOTHEICAL PROTEIN M	1.55e+03
898	4	21.1	1060	1	YKRI_CAEEL	HYPOTHEICAL 118.3 KD	1.55e+03	971	4	21.1	1761	1	YKRI_CAEEL	PUTATIVE ABC TRANSPORT	1.55e+03
899	4	21.1	1062	1	YKRI_CAEEL	PUTATIVE 118.2 KD TRAN	1.55e+03	972	4	21.1	1788	1	YKRI_CAEEL	HYPOTHEICAL 200.6 KD	1.55e+03

```

973 4 21.1 1799 1 Y025_CABEL 1.55e+03
974 4 21.1 1807 1 VT22_XENLA 1.55e+03
975 4 21.1 1888 1 YD72_SCHPO 1.55e+03
976 4 21.1 1895 1 GLS2_YEAST 1.55e+03
977 4 21.1 1972 1 MYST_RABIT 1.55e+03
978 4 21.1 1984 1 YLDRONE 1.55e+03
979 4 21.1 1995 1 YC7_KHLRE 1.55e+03
980 4 21.1 2014 1 YJ07_YEAST 1.55e+03
981 4 21.1 2052 1 YD08_SCHPO 1.55e+03
982 4 21.1 2054 1 YCF2_PINTH 1.55e+03
983 4 21.1 2131 1 YCF2_SPIOL 1.55e+03
984 4 21.1 2136 1 YCF2_WARPO 1.55e+03
985 4 21.1 2151 1 RRPL_RANTV 1.55e+03
986 4 21.1 2165 1 RRPL_HRSVA 1.55e+03
987 4 21.1 2185 1 PYR1_DICDI 1.55e+03
988 4 21.1 2194 1 SC16_YEAST 1.55e+03
989 4 21.1 2216 1 YCF2_EPIVI 1.55e+03
990 4 21.1 2242 1 PYR1_SQUAC 1.55e+03
991 4 21.1 2280 1 YCF2_TOBAC 1.55e+03
992 4 21.1 2334 1 WAPA_BACSU 1.55e+03
993 4 21.1 2339 1 RPC1_PLAFA 1.55e+03
994 4 21.1 2366 1 TOXB_GLODI 1.55e+03
995 4 21.1 2376 1 YIN9_YEAST 1.55e+03
996 4 21.1 2491 1 MPRI_HUMAN 1.55e+03
997 4 21.1 2499 1 MPRI_BOVIN 1.55e+03
998 4 21.1 2763 1 TEGU_VZVD 1.55e+03
999 4 21.1 3421 1 TEGU_HSVB 1.55e+03
1000 4 21.1 4036 1 RRPL_DUGBV 1.55e+03

```

## ALIGNMENTS

```

RESULT 1
ID YK2_YEAST STANDARD; PRT; 151 AA.
AC P36108;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL 16.7 KD PROTEIN MRP17-MET14 INTERGENIC REGION.
GN YK1002W.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACHAROMYCES.
RN [1]
RA BOYER J., PASCOLO S., RICHARD G.F., GHAVINI M., COLLEAUX L.,
RA THIERRY A., MONNIER A.L., DUJON B.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RA DUESTERHOEFT A., MÖBSTL D., POEHLMANN R., PHILIPPSEN P.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 228002; G485989; -
DR PIR; S37812; S37812.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 151 AA; 16679 MW; 18CFDADF CRC32;

```

Query Match 31.6%; Score 6; DB 1; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 1.27e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 11 RIQAVR 16  
 |||||

```

QY 13 RIQAVR 18
RESULT 2
ID HS49_YEAST STANDARD; PRT; 213 AA.
AC Q99181;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HSH49 PROTEIN.
GN HSH49 OR YOR319W OR O6142.
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RA MEDLINE; 97051589;
RA PEARSON B.M., HERNANDO Y., PAYNE J., WOLF S.S., KALOGEROPOULOS A.,
RA SCHWEIZER M.;
RT "Sequencing of a 35.71 kb DNA segment on the right arm of yeast
RT chromosome XV reveals regions of similarity to chromosomes I and
RT XIII";
RL YEAST 12:1021-1031(1996).
CC -1- FUNCTION: POSSIBLE SF3B-LIKE FACTOR.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: TO MAMMALIAN SAP 49 AND TO C.ELEGANS C08B11.5.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RNP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z75227; E252149; -
DR EMBL; X90565; G940850; -
DR SGD; L0003014; HSH49.
DR PROSITE; PS00030; RNP_1; FALSE_NEG.
DR PFAM; PF00076; rrm; 2.
DR HSP; P19339; 2SXL.
DR KW MRNA PROCESSING; RNA-BINDING; NUCLEAR PROTEIN; REPEAT.
FT DOMAIN 11 16 RNA-BINDING (RNP2) (BY SIMILARITY).
FT DOMAIN 50 57 RNA-BINDING (RNP1) (BY SIMILARITY).
FT DOMAIN 110 115 RNA-BINDING (RNP2) (BY SIMILARITY).
FT DOMAIN 149 155 RNA-BINDING (RNP1) (BY SIMILARITY).
SQ SEQUENCE 213 AA; 24503 MW; 0356C013 CRC32;

```

Query Match 31.6%; Score 6; DB 1; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 1.27e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 141 FYLSNG 146  
 |||||  
 QY 7 FYLSNG 12

```

RESULT 3
ID FL3H_CALCH STANDARD; PRT; 356 AA.
AC Q05963;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE NARINGENIN, 2-OXOGLUTARATE 3-DIOXYGENASE (EC 1.14.11.9) (FLAVONONE-
DE 3-HYDROXYLASE) (FHT).
GN FHT.
OS CALLISTEPHUS CHINENSIS (CHINA ASTER).
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; TRACHEOPHYTA;
OC EUPHYLLIPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS;
OC ASTERIDAE; ASTERALES; ASTERACEAE; CALLISTEPHUS.
RN [1]

```



DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR	PFAM; PF00005; ABC_tran; 1.
KW	HYPOTHETICAL PROTEIN; ATP-BINDING; TRANSPORT; TRANSMEMBRANE.
FT	TRANSMEM 16 36 POTENTIAL.
FT	TRANSMEM 48 68 POTENTIAL.
FT	TRANSMEM 125 145 POTENTIAL.
FT	TRANSMEM 149 169 POTENTIAL.
FT	TRANSMEM 240 260 POTENTIAL.
FT	TRANSMEM 270 290 POTENTIAL.
FT	NP_BIND 356 363 ATP (POTENTIAL).
SEQ	SEQUENCE 547 AA; 61552 MW; 575960F4 CRC32;
Query Match	
Best Local Similarity 100.0%; Pred. No.1.27e+00;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	440 LSNMRI 445
QY	
	9 LSNMRI 14
RESULT	7
ID	IP44_SHIFL STANDARD; PRT; 574 AA.
AC	FL8009;
DT	01-NOV-1990 (REL. 16, CREATED)
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE	65.4 KD ANTIGEN.
GN	IP4H 4.5.
OS	SHIGELLA FLEXNERI.
OG	PLASMID 210 KB INVASION PWR100.
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC	SHIGELLA.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=M90T / SEROTYPE 5;
RX	MEDLINE; 92167809.
RA	VENKATESAN M.M., BUYSE J.M., HARTMAN A.B.;
RT	"Sequence variation in two ipah genes of Shigella flexneri 5 and
RT	homology to the LRG-like family of proteins.;"
RL	MOL. MICROBIOL. 5:2435-2445(1991).
RN	[2]
RP	SEQUENCE OF 1-208 FROM N.A.
RC	STRAIN=M90T / SEROTYPE 5;
RX	MEDLINE; 90202708
RA	HARTMAN A.B., VENKATESAN M.M., OAKS E.V., BUYSE J.M.;
RT	"Sequence and molecular characterization of a multicopy invasion
RT	plasmid antigen gene, ipah, of Shigella flexneri.;"
RL	J. BACTERIOL. 172:1905-1915(1990).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; M76445; G152752; -
DR	EMBL; M32063; G152748; -
DR	PIR; B35149; B35149.
DR	PFAM; PF00560; LRR: 4.
KW	ANTIGEN; PLASMID; VIRULENCE; REPEAT; MULTIGENE FAMILY.
FT	DOMAIN 74 228
FT	REPEAT 74 87
FT	REPEAT 94 107
FT	REPEAT 114 127
FT	REPEAT 135 148
FT	REPEAT 155 168
FT	REPEAT 175 188
FT	REPEAT 195 208
FT	REPEAT 215 228
FT	REPEAT 235 248
FT	REPEAT 255 268
FT	REPEAT 275 288
FT	REPEAT 295 308
FT	REPEAT 315 328
FT	REPEAT 335 348
FT	REPEAT 355 368
FT	REPEAT 375 388
FT	REPEAT 395 408
FT	REPEAT 415 428
FT	REPEAT 435 448
FT	REPEAT 455 468
FT	REPEAT 475 488
FT	REPEAT 495 508
FT	REPEAT 515 528
FT	REPEAT 535 548
FT	REPEAT 555 568
FT	REPEAT 575 588
FT	REPEAT 595 608
FT	REPEAT 615 628
FT	REPEAT 635 648
FT	REPEAT 655 668
FT	REPEAT 675 688
FT	REPEAT 695 708
FT	REPEAT 715 728
FT	REPEAT 735 748
FT	REPEAT 755 768
FT	REPEAT 775 788
FT	REPEAT 795 808
FT	REPEAT 815 828
FT	REPEAT 835 848
FT	REPEAT 855 868
FT	REPEAT 875 888
FT	REPEAT 895 908
FT	REPEAT 915 928
FT	REPEAT 935 948
FT	REPEAT 955 968
FT	REPEAT 975 988
FT	REPEAT 995 1008
FT	REPEAT 1015 1028
FT	REPEAT 1035 1048
FT	REPEAT 1055 1068
FT	REPEAT 1075 1088
FT	REPEAT 1095 1108
FT	REPEAT 1115 1128
FT	REPEAT 1135 1148
FT	REPEAT 1155 1168
FT	REPEAT 1175 1188
FT	REPEAT 1195 1208
FT	REPEAT 1215 1228
FT	REPEAT 1235 1248
FT	REPEAT 1255 1268
FT	REPEAT 1275 1288
FT	REPEAT 1295 1308
FT	REPEAT 1315 1328
FT	REPEAT 1335 1348
FT	REPEAT 1355 1368
FT	REPEAT 1375 1388
FT	REPEAT 1395 1408
FT	REPEAT 1415 1428
FT	REPEAT 1435 1448
FT	REPEAT 1455 1468
FT	REPEAT 1475 1488
FT	REPEAT 1495 1508
FT	REPEAT 1515 1528
FT	REPEAT 1535 1548
FT	REPEAT 1555 1568
FT	REPEAT 1575 1588
FT	REPEAT 1595 1608
FT	REPEAT 1615 1628
FT	REPEAT 1635 1648
FT	REPEAT 1655 1668
FT	REPEAT 1675 1688
FT	REPEAT 1695 1708
FT	REPEAT 1715 1728
FT	REPEAT 1735 1748
FT	REPEAT 1755 1768
FT	REPEAT 1775 1788
FT	REPEAT 1795 1808
FT	REPEAT 1815 1828
FT	REPEAT 1835 1848
FT	REPEAT 1855 1868

```
SQ SEQUENCE 574 AA; 65408 MW; 1C2DE575 CRC32;

Query Match 31.6%; Score 6; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 1.27e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 49 RIOAVR 54
   |||||
QY 13 RIOAVR 18

RESULT 8
ID VGL2.CVCAI STANDARD; PRT; 1451 AA.
AC P36300.
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN).
GN S.
OS CANINE ENTERIC CORONAVIRUS (STRAIN INSNVC-1) (CCV).
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES; NO DNA STAGE; NIDOVIRALES;
OC CORONAVIRIDAE; CORONAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX HORSBURGH B.C., BRIERLEY I., BROWN T.D.K.;
RA "Analysis of a 9.6 kb sequence from the 3' end of canine coronavirus
RT genomic RNA.";
RL J. GEN. VIROL. 73:2849-2862(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D13096; G406195; -.
CC DR EMBL: A22884; E185326; -.
CC DR PIR: JQ1719; JQ1719.
CC KW GLYCOPROTEIN; ENVELOPE PROTEIN; TRANSMEMBRANE; SIGNAL.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 1451 SPIKE E2 GLYCOPROTEIN.
FT DOMAIN 19 1393 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1394 1412 POTENTIAL.
FT DOMAIN 1413 1451 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1413 1434 CYS-RICH.
FT CARBOHYD 28 28 POTENTIAL.
FT CARBOHYD 66 66 POTENTIAL.
FT CARBOHYD 94 94 POTENTIAL.
FT CARBOHYD 142 142 POTENTIAL.
FT CARBOHYD 175 175 POTENTIAL.
FT CARBOHYD 209 209 POTENTIAL.
FT CARBOHYD 235 235 POTENTIAL.
FT CARBOHYD 242 242 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CARBOHYD 338 338 POTENTIAL.
FT CARBOHYD 349 349 POTENTIAL.
FT CARBOHYD 366 366 POTENTIAL.
FT CARBOHYD 379 379 POTENTIAL.
FT CARBOHYD 409 409 POTENTIAL.
FT CARBOHYD 453 453 POTENTIAL.
FT CARBOHYD 520 520 POTENTIAL.
FT CARBOHYD 536 536 POTENTIAL.
FT CARBOHYD 557 557 POTENTIAL.
FT CARBOHYD 707 707 POTENTIAL.
FT CARBOHYD 728 728 POTENTIAL.
FT CARBOHYD 783 783 POTENTIAL.
FT CARBOHYD 821 821 POTENTIAL.
FT CARBOHYD 836 836 POTENTIAL.
FT CARBOHYD 842 842 POTENTIAL.
FT CARBOHYD 923 923 POTENTIAL.

SQ SEQUENCE 1076 1076 POTENTIAL.
FT CARBOHYD 1202 1202 POTENTIAL.
FT CARBOHYD 1296 1296 POTENTIAL.
FT CARBOHYD 1313 1313 POTENTIAL.
FT CARBOHYD 1326 1326 POTENTIAL.
FT CARBOHYD 1343 1343 POTENTIAL.
FT CARBOHYD 1360 1360 POTENTIAL.
FT CARBOHYD 1373 1373 POTENTIAL.
SQ SEQUENCE 1451 AA; 160467 MW; 9B36BEEA CRC32;

Query Match 31.6%; Score 6; DB 1; Length 1451;
Best Local Similarity 100.0%; Pred. No. 1.27e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1263 DEXFYL 1268
   |||||
QY 4 DEXFYL 9

RESULT 9
ID RPAL_RAT STANDARD; PRT; 1716 AA.
AC O54889;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-DIRECTED RNA POLYMERASE I LARGEST SUBUNIT (EC 2.7.7.6) (RNA
DE POLYMERASE I 194 KD SUBUNIT) (RPAL94).
GN RPOL-4 OR RPAL.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98086309.
RA HANNAH R.D., HEMPEL W.M., CAVANAUGH A., ARINO T., DIMITROV S.I.,
RA MOSS T., ROTHBLUM L.;
RT "Affinity purification of mammalian RNA polymerase I. Identification
of an associated kinase.";
RL J. BIOL. CHEM. 273:1257-1267(1998).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES. RNA POLYMERASE I IS ESSENTIALLY USED TO TRANSCRIBE
CC RIBOSOMAL DNA UNITS.
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N PYROPHOSPHATE +
CC RNA(N).
CC -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE LARGEST COMPONENT OF
CC RNA POLYMERASE I.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN
CC EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
CC POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
CC AND TRNA GENES.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF025425; G2739050; -.
CC DR TRANSFERASE; DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC;
CC ZINC-FINGER; NUCLEAR PROTEIN. C2H2-TYPE (POTENTIAL).
FT ZN_FING 64 80
SQ SEQUENCE 1716 AA; 194191 MW; DA211E87 CRC32;

Query Match 31.6%; Score 6; DB 1; Length 1716;
Best Local Similarity 100.0%; Pred. No. 1.27e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db 1497 RIOAVR 1502
QY 13 RIOAVR 18

RESULT 10
ID RPA1_MOUSE STANDARD; PRT: 1717 AA.
AC Q35134;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA-DIRECTED RNA POLYMERASE I LARGEST SUBUNIT (EC 2.7.7.6) (RNA
POL-4 OR RPA1.
GN RPA1-4 OR RPA1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97379976.
RA SEITHER P., COY J.F., POUSKA A., GRUMMT I.;
RT "Molecular cloning and characterization of the cDNA encoding the
largest subunit of mouse RNA polymerase I.";
RL MOL. GEN. GENET. 255:180-186(1997).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES. RNA POLYMERASE I IS ESSENTIALLY USED TO TRANSCRIBE
RIBOSOMAL DNA UNITS.
CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
RNA(N).
CC -!- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14
DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE LARGEST COMPONENT OF
RNA POLYMERASE I.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE FOUND IN
EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA PRECURSOR,
POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE III FOR 5S
AND TRNA GENES.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF000938; G2330007; -.
CC MGD: MGI:1096397; RPA1-4.
CC PFAM: PF00623; RNA_POL_A; 1.
CC TRANSFERASE; DNA-DIRECTED RNA POLYMERASE; TRANSCRIPTION; ZINC;
CC ZINC-FINGER; NUCLEAR PROTEIN.
CC ZN_FING 64 80 C2H2-TYPE (POTENTIAL).
CC SEQUENCE 1717 AA; 194079 MW; A3F355F3 CRC32;

Query Match 31.6%; Score 6; DB 1; Length 1717;
Best Local Similarity 100.0%; Pred. No. 1.27e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1498 RIOAVR 1503
QY 13 RIOAVR 18

RESULT 11
ID YC2X_GUITH STANDARD; PRT: 65 AA.
AC O78421;
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 8.1 KD PROTEIN (ORF65).
OS GUILLARDIA THETA (CRYPTOMONAS PHI).

```

---

```

OG CHLOROPLAST.
OC EUKARYOTA; CRYPTOPHYTA; CRYPTOMONADACEAE; GUILLARDIA.
RN [1]
RP SEQUENCE FROM N.A.
RA DOUGLAS S.E., PENNY S.L.;
RT "The plastid genome from the cryptomonad alga, Guillardia theta:
complete sequence and conserved syntenic groups confirm its common
ancestry with red algae.";
RL J. MOL. EVOL. 0:0-0(1998).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: AF041468; G3602945; -.
CC HYPOTHETICAL PROTEIN; CHLOROPLAST.
CC SEQUENCE 65 AA; 8086 MW; FCB4FF6D CRC32;

Query Match 26.3%; Score 5; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 5.88e+01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 47 EK FYL 51
QY 5 EK FYL 9

RESULT 12
ID PSPB_PIG STANDARD; PRT: 79 AA.
AC P15782;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE PULMONARY SURFACTANT-ASSOCIATED PROTEIN B (SP-B) (8 KD PROTEIN)
(PULMONARY SURFACTANT-ASSOCIATED PROTEOLIPID SPL(PHE)).
GN SFTPB OR SFTP3.
OS SUS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE.
RX MEDLINE: 88166729.
RA CURSTEDT T., JOHANSSON J., BARROS-SOEDERLING J., ROBERTSON B.,
NILSSON G., WESTBERG M., JOERNVALL H.;
RT "Low-molecular-mass surfactant protein type 1. The primary structure
of a hydrophobic 8-kDa polypeptide with eight half-cysteine
residues.";
RL EUR. J. BIOCHEM. 172:521-525(1988).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE: 91299745.
RA JOHANSSON J., CURSTEDT T., JOERNVALL H.;
RT "Surfactant protein B: disulfide bridges, structural properties, and
kringle similarities.";
RL BIOCHEMISTRY 30:6917-6921(1991).
CC -!- FUNCTION: PULMONARY SURFACTANT ASSOCIATED PROTEINS PROMOTE
ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
PER METER.
CC -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE
ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-
BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC
PROTEINS (SP-B AND SP-C).
CC PIR: S00363; LNPG1.
CC SURFACE FILM; GASEOUS EXCHANGE.
CC DISULFID 8 77

```

```
FT DISULFID 11 71
FT DISULFID 35 46
FT DISULFID 48 48
FT VARIANT 57 57
SQ SEQUENCE 79 AA: 8714 MW: 2BC88021 CRC32:
  Query Match 26.3%; Score 5; DB 1; Length 79;
  Best Local Similarity 100.0%; Pred. No. 5.88e+01;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 17 RIQAV 21
  |||||
QY 13 RIQAV 17

RESULT 13
ID VC14_VACCC STANDARD; PRT; 82 AA.
AC P21045;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE PROTEIN C14.
GN C14L.
OS VACCINIA VIRUS (STRAIN COPENHAGEN).
OC VIRUSES: DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
OC ORTHOPOXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91021027.
RA GOBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL "The complete DNA sequence of vaccinia virus.";
RL VIROLOGY 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA GOBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,
RA PAOLETTI E.;
RL VIROLOGY 179:517-563(1990).

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M35027; G335330;
DR PIR: I42502; I42502.
SQ SEQUENCE 82 AA: 9304 MW: 57FA451D CRC32:

  Query Match 26.3%; Score 5; DB 1; Length 82;
  Best Local Similarity 100.0%; Pred. No. 5.88e+01;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 64 YLSNG 68
  |||||
QY 8 YLSNG 12

RESULT 14
ID YIDD_HAEIN STANDARD; PRT; 86 AA.
AC P44972;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 9.5 KD PROTEIN IN RNPA 3' REGION.
GN H11000.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=RD / KW20;
RX MEDLINE: 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
CC -!- SIMILARITY: TO THE CORRESPONDING ORF IN RELATED BACTERIA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U32781; G1574031;
DR TIGR: H11000;
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 86 AA: 9564 MW: F38629AE CRC32:

  Query Match 26.3%; Score 5; DB 1; Length 86;
  Best Local Similarity 100.0%; Pred. No. 5.88e+01;
  Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 81 NNDEK 85
  |||||
QY 2 NNDEK 6

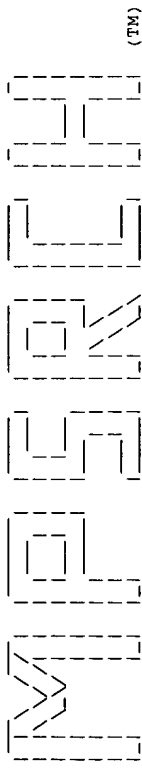
RESULT 15
ID RR14_PEA STANDARD; PRT; 100 AA.
AC P05638;
DT 01-NOV-1988 (REL. 09, CREATED)
DT 01-NOV-1988 (REL. 09, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S14.
GN RPS14.
OS PISUM SATIVUM (GARDEN PEA).
OC CHLOROPLAST.
OC EUKARYOTA; VIRIDIPHYTES; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC FABALES; FABACEAE; PAPILIONOIDEAE; PISUM.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87203396.
RA LEHMBECK J., STUMMANN B.M., HENNINGSSEN K.W.;
RT "Sequence of two regions of pea chloroplast DNA, one with the genes
RT rps14, trnM and trnG-GCC, and one with the genes trnP-UGG and
RT trnW-CCA.";
RL NUCLEIC ACIDS RES. 15:3630-3630(1987).
CC -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X05394; G12174;
DR PIR: A27523; A27523.
DR PROSITE: PS00527; RIBOSOMAL_S14; 1.
DR PFAM: PF00253; S14; 1.
```

DR MENDEL; 4142; PISsa; rps14; 1.  
KW RIBOSOMAL PROTEIN; CHLOROPLAST.  
SQ SEQUENCE 100 AA; 11608 MW; 39D01141 CRC32;  
Query Match 26.3%; Score 5; DB 1; Length 100;  
Best Local Similarity 100.0%; Pred. No. 5.88e-01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 18 EKFYL 22  
|||||  
QY 5 EKFYL 9

Search completed: Sat Aug 28 14:37:48 1999  
Job time : 44 secs.



\*\*\*\*\*



\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat Aug 28 14:38:06 1999; MasPar time 7.21 Seconds  
143.794 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-049-696-42  
Description: (1-19) from US09049696.pep  
Perfect Score: 19  
Sequence: 1 YNDEKEYLSNGRIQAVRC 19

Scoring table: TABLE uniprotatable  
Gap 60

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 1000 summaries

Database: spiremb19  
1:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phase 10:sp\_plant 11:sp\_rodent 12:sp\_unclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 2.433; Variance 0.358; scale 6.787

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description	Pred. No.
1	10	52.6	913	11	O88826	3.73e-09
2	7	36.8	433	5	GOB-5 PROTEIN.	1.17e-02
3	6	31.6	97	1	F39B2.8 PROTEIN.	1.02e+00
4	6	31.6	170	5	HYPOTHETICAL 10.4 KD P	1.02e+00
5	6	31.6	181	2	C54F6.14 PROTEIN.	1.02e+00
6	6	31.6	314	2	C12_ORF1810 PROTEIN.	1.02e+00
7	6	31.6	316	2	AUXIN-INDUCED PROTEIN.	1.02e+00
8	6	31.6	316	2	SIMILAR TO ISOMALTASE-	1.02e+00
9	6	31.6	332	10	XYLOGLUCAN ENDOTRANSGL	1.02e+00
10	6	31.6	372	10	FLAVANONE 3-HYDROXYLAS	1.02e+00
11	6	31.6	374	10	FLAVANONE 3-HYDROXYLAS	1.02e+00
12	6	31.6	382	5	FLAVANONE 3-DIOXYGENA	1.02e+00
13	6	31.6	410	2	LTR-RETROTRANSPOSON SK	1.02e+00
14	6	31.6	410	2	TECA4.	1.02e+00
15	6	31.6	458	1	CHLOROBENZENE DIOXYGEN	1.02e+00
16	6	31.6	545	2	CONSERVED PROTEIN.	1.02e+00
17	6	31.6	621	10	ADENINE SPECIFIC DNA M	1.02e+00
18	6	31.6	734	3	FLAVANONE 3-HYDROXYLAS	1.02e+00
19	6	31.6	848	5	HYPOTHETICAL 85.3 KD P	1.02e+00
20	6	31.6	919	5	M03C11.2 PROTEIN.	1.02e+00
					HYPOTHETICAL 103.3 KD	1.02e+00

GLUTAMATE RECEPTOR SUB 1.02e+00  
GLUTAMATE RECEPTOR CHA 1.02e+00  
GLUTAMATE RECEPTOR DEL 1.02e+00  
SMDR2. 1.02e+00  
GLUCOSYLTRANSFERASE (E 1.02e+00  
F31C3.3 PROTEIN. 6.00e+01  
MHC-DQA-1\*03 (FRAGMENT 6.00e+01  
MHC CLASS II DQ ALPHA 6.00e+01  
MHC-DQ-ALPHA 1\*02 (FRA 6.00e+01  
MHC ALPHA CHAIN (FRAGM 6.00e+01  
ORF 2. 6.00e+01  
HYPOTHETICAL 11.4 KD P 6.00e+01  
ORF103. 6.00e+01  
108AA LONG HYPOTHETICA 6.00e+01  
POLO-LIKE KINASE ISOFO 6.00e+01  
SIMILAR TO SINGLE STRA 6.00e+01  
KERATIN TYPE II CYTOSK 6.00e+01  
HYPOTHETICAL 14.5 KD P 6.00e+01  
ALLENE OXIDE SYNTHASE. 6.00e+01  
1408A LONG HYPOTHETICA 6.00e+01  
H-ATPASE HOMOLOG (FRAG 6.00e+01  
HYPOTHETICAL 16.2 KD P 6.00e+01  
HYPOTHETICAL 16.2 KD P 6.00e+01  
ENVELOPE GLYCOPROTEIN 6.00e+01  
F17A2.4 PROTEIN. 6.00e+01  
M PROTEIN (FRAGMENT). 6.00e+01  
KDPD. 6.00e+01  
HYPOTHETICAL 18.1 KD P 6.00e+01  
ENVELOPE PROTEIN (FRAG 6.00e+01  
MOBILIZATION PROTEIN. 6.00e+01  
MBEB. 6.00e+01  
1748A LONG HYPOTHETICA 6.00e+01  
CAPSID PROTEIN. 6.00e+01  
FERRIC UPTAKE REGULATI 6.00e+01  
ORF177 (FRAGMENT). 6.00e+01  
ENVELOPE GLYCOPROTEIN 6.00e+01  
PREPROENDOTHELIN-1 (FR 6.00e+01  
PUTATIVE LIPOPROTEIN. 6.00e+01  
FIX23-4. 6.00e+01  
TRANSCRIPTIONAL REGULA 6.00e+01  
B21R PROTEIN. 6.00e+01  
ORF C04037. 6.00e+01  
TMBC (FRAGMENT). 6.00e+01  
ACHAETE. 6.00e+01  
ACHAETE GENE. 6.00e+01  
3-ISOPROPYLMALATE DEHY 6.00e+01  
AGTHR4 GENE AND PUTATI 6.00e+01  
SERINE/THREONINE PROTE 6.00e+01  
BSCL. 6.00e+01  
N-METHYL-D-ASPARTATE R 6.00e+01  
PUTATIVE TWO-COMPONENT 6.00e+01  
CARR. 6.00e+01  
T26B15.11 PROTEIN. 6.00e+01  
HYPOTHETICAL 27.4 KD P 6.00e+01  
DNA CHROMOSOME 4, ESSA 6.00e+01  
SIGNAL-TRANSDUCING PRO 6.00e+01  
PUTATIVE PEPTIDASE. 6.00e+01  
COSMID K08D10. 6.00e+01  
SIMILARITY TO COLD-SHO 6.00e+01  
267AA LONG HYPOTHETICA 6.00e+01  
COSMID R04A9. 6.00e+01  
SIMILARITY TO A COLD-S 6.00e+01  
SERINE PROTEASE. 6.00e+01  
LATE NODULIN NVF-28 PR 6.00e+01  
HYPOTHETICAL 31.7 KD P 6.00e+01  
HYPOTHETICAL 32.3 KD P 6.00e+01  
NODULIN-30 (N-30). 6.00e+01  
F28D9.1 PROTEIN. 6.00e+01  
LATE NODULIN NVF-30B P 6.00e+01  
PROTEIN WITH PARTIAL H 6.00e+01  
(HHV-6). 6.00e+01  
DNA SIMILAR TO P450 GE 6.00e+01  
MATRIX PROTEIN. 6.00e+01

21 6 31.6 927 11 063225  
22 6 31.6 1009 11 061627  
23 6 31.6 1009 11 062640  
24 6 31.6 1254 5 026599  
25 6 31.6 1508 2 052224  
26 6 31.6 2025 5 062191  
27 5 26.3 62 7 030995  
28 5 26.3 62 7 019302  
29 5 26.3 63 7 030337  
30 5 26.3 82 7 031291  
31 5 26.3 93 2 052151  
32 5 26.3 99 2 032775  
33 5 26.3 103 8 020116  
34 5 26.3 108 1 059087  
35 5 26.3 111 11 088432  
36 5 26.3 113 2 094590  
37 5 26.3 116 6 029386  
38 5 26.3 129 2 073512  
39 5 26.3 132 1 029737  
40 5 26.3 140 1 057888  
41 5 26.3 141 2 006689  
42 5 26.3 142 2 073106  
43 5 26.3 147 14 009497  
44 5 26.3 148 14 040184  
45 5 26.3 149 5 019502  
46 5 26.3 153 2 085811  
47 5 26.3 156 2 077876  
48 5 26.3 164 2 006572  
49 5 26.3 167 14 080718  
50 5 26.3 172 2 052254  
51 5 26.3 172 2 069425  
52 5 26.3 174 1 058019  
53 5 26.3 175 14 083946  
54 5 26.3 176 2 051592  
55 5 26.3 177 2 054754  
56 5 26.3 178 14 091871  
57 5 26.3 180 11 097338  
58 5 26.3 182 2 053569  
59 5 26.3 188 2 052933  
60 5 26.3 188 2 059802  
61 5 26.3 190 14 072758  
62 5 26.3 192 1 095975  
63 5 26.3 195 2 051991  
64 5 26.3 201 5 077026  
65 5 26.3 201 5 024639  
66 5 26.3 205 2 031294  
67 5 26.3 209 3 000064  
68 5 26.3 211 10 093770  
69 5 26.3 212 2 068538  
70 5 26.3 221 13 091414  
71 5 26.3 222 2 086632  
72 5 26.3 244 2 051932  
73 5 26.3 248 10 080892  
74 5 26.3 250 2 034684  
75 5 26.3 251 10 023458  
76 5 26.3 254 2 024971  
77 5 26.3 257 2 034493  
78 5 26.3 259 5 021317  
79 5 26.3 265 5 091398  
80 5 26.3 267 1 058931  
81 5 26.3 267 5 021692  
82 5 26.3 267 5 091306  
83 5 26.3 268 5 017030  
84 5 26.3 269 10 007514  
85 5 26.3 277 3 013951  
86 5 26.3 278 2 051287  
87 5 26.3 279 10 007515  
88 5 26.3 279 5 062185  
89 5 26.3 281 10 024537  
90 5 26.3 285 14 089432  
91 5 26.3 287 10 041665  
92 5 26.3 294 3 013469  
93 5 26.3 295 14 083325

94	5	26.3	295	2	054416	MALONATE DECARBOXYLASE	6.00e+01	167	5	26.3	366	2	088001	PUTATIVE AMINO-SUGAR B	6.00e+01
95	5	26.3	295	10	Q07524	ENDOGLUCANASE 1 (EC 3.	6.00e+01	168	5	26.3	367	10	Q96472	FLAVANONE 3-HYDROXYLAS	6.00e+01
96	5	26.3	297	10	Q07516	NODULIN-32B (N-32B) (F	6.00e+01	169	5	26.3	367	10	Q42673	PAPAYA PROTEINASE OMEG	6.00e+01
97	5	26.3	298	10	Q41667	PROTEIN SHOWING PARTIA	6.00e+01	170	5	26.3	368	2	Q06994	HYPOETHICAL 42.2 KD P	6.00e+01
98	5	26.3	300	14	Q83324	MATRIX PROTEIN.	6.00e+01	171	5	26.3	369	5	Q18033	T06G6.7 PROTEIN.	6.00e+01
99	5	26.3	301	5	Q44817	F52C6.10 PROTEIN.	6.00e+01	172	5	26.3	369	10	Q40754	FLAVANONE 3-HYDROXYLAS	6.00e+01
100	5	26.3	310	6	Q28810	KERATIN 8 (FRAGMENT).	6.00e+01	173	5	26.3	372	10	Q43262	FLAVANONE 3-BETA-HYDRO	6.00e+01
101	5	26.3	308	5	Q09926	K08F8.5B PROTEIN.	6.00e+01	174	5	26.3	374	5	Q16979	T03D3.5 PROTEIN.	6.00e+01
102	5	26.3	312	14	Q69551	U16EXON1-2.	6.00e+01	175	5	26.3	374	5	Q19973	CODER FOR BY C. ELEGAN	6.00e+01
103	5	26.3	316	10	Q81606	1-AMINOCYCLOPROPANE-1-	6.00e+01	176	5	26.3	375	2	Q69189	DNA GYRASE B (EC 5.99.	6.00e+01
104	5	26.3	320	5	Q46102	EG:8D8.8 PROTEIN.	6.00e+01	177	5	26.3	375	6	Q77802	ANGIOPOIETIN-2 (FRASME	6.00e+01
105	5	26.3	320	5	Q18887	SIMILARITY TO FREQUENI	6.00e+01	178	5	26.3	379	10	Q80417	NTC12 PROTEIN.	6.00e+01
106	5	26.3	321	1	Q50744	F420-DEPENDENT METHYLE	6.00e+01	179	5	26.3	379	10	Q80385	GIBBERELLIN 20-OXIDASE	6.00e+01
107	5	26.3	321	1	Q27784	COENZYME F420-DEPENDEN	6.00e+01	180	5	26.3	380	10	Q93473	GIBBERELLIN 20-OXIDASE	6.00e+01
108	5	26.3	321	1	Q50540	COENZYME F420-DEPENDEN	6.00e+01	181	5	26.3	380	10	Q42277	GA 20-OXIDASE.	6.00e+01
109	5	26.3	323	10	Q24536	NODULIN 28/32 PRECURSO	6.00e+01	182	5	26.3	381	10	Q04280	GIBBERELLIN 20-OXIDASE	6.00e+01
110	5	26.3	327	5	Q18955	D1081.2 PROTEIN.	6.00e+01	183	5	26.3	382	14	Q83907	DNA BINDING PROTEIN.	6.00e+01
111	5	26.3	328	2	Q57527	HYPOTHETICAL PROTEIN H	6.00e+01	184	5	26.3	382	2	Q94791	LYCOPENE CYCLASE.	6.00e+01
112	5	26.3	331	2	Q24994	BETA-KETOACYL-ACYL CAR	6.00e+01	185	5	26.3	382	2	Q85764	PUTATIVE FMN2-DEPENDE	6.00e+01
113	5	26.3	335	14	Q93193	MEMBRANE ASSOCIATED PR	6.00e+01	186	5	26.3	383	2	Q56181	SENSOR KINASE PHOR.	6.00e+01
114	5	26.3	335	14	Q89260	MEMBRANE ASSOCIATED PR	6.00e+01	187	5	26.3	383	2	Q31150	ALCOHOL DEHYDROGENASE	6.00e+01
115	5	26.3	335	14	Q86498	MATRIX PROTEIN.	6.00e+01	188	5	26.3	385	10	Q04281	GIBBERELLIN 20-OXIDASE	6.00e+01
116	5	26.3	335	14	Q89120	M PROTEIN.	6.00e+01	189	5	26.3	387	2	Q95416	NIRJ.	6.00e+01
117	5	26.3	335	14	Q89121	M PROTEIN.	6.00e+01	190	5	26.3	389	5	Q94811	ANTIGEN.	6.00e+01
118	5	26.3	335	14	Q89326	MATRIX PROTEIN M.	6.00e+01	191	5	26.3	389	14	Q11528	ENVELOPE GLYCOPROTEIN	6.00e+01
119	5	26.3	335	14	Q83719	MATRIX PROTEIN.	6.00e+01	192	5	26.3	390	14	Q11530	ENVELOPE GLYCOPROTEIN	6.00e+01
120	5	26.3	335	14	Q04254	MATRIX PROTEIN.	6.00e+01	193	5	26.3	390	14	Q11539	ENVELOPE GLYCOPROTEIN	6.00e+01
121	5	26.3	335	14	Q09111	MATRIX PROTEIN.	6.00e+01	194	5	26.3	390	14	Q11533	ENVELOPE GLYCOPROTEIN	6.00e+01
122	5	26.3	335	14	Q83624	MATRIX PROTEIN.	6.00e+01	195	5	26.3	390	14	Q11535	ENVELOPE GLYCOPROTEIN	6.00e+01
123	5	26.3	335	14	Q04253	MATRIX PROTEIN.	6.00e+01	196	5	26.3	390	14	Q11534	ENVELOPE GLYCOPROTEIN	6.00e+01
124	5	26.3	335	14	Q04256	MATRIX PROTEIN.	6.00e+01	197	5	26.3	390	14	Q11527	ENVELOPE GLYCOPROTEIN	6.00e+01
125	5	26.3	335	14	Q83722	MATRIX PROTEIN.	6.00e+01	198	5	26.3	390	14	Q11531	ENVELOPE GLYCOPROTEIN	6.00e+01
126	5	26.3	335	14	Q83728	MATRIX PROTEIN.	6.00e+01	199	5	26.3	390	14	Q11531	ENVELOPE GLYCOPROTEIN	6.00e+01
127	5	26.3	335	14	Q83720	MATRIX PROTEIN.	6.00e+01	200	5	26.3	390	14	Q11536	ENVELOPE GLYCOPROTEIN	6.00e+01
128	5	26.3	335	14	Q83712	MATRIX PROTEIN.	6.00e+01	201	5	26.3	393	2	Q95550	ORF393 PROTEIN.	6.00e+01
129	5	26.3	335	14	Q83716	MATRIX PROTEIN.	6.00e+01	202	5	26.3	399	4	Q13167	DOPAMINE D3 RECEPTOR.	6.00e+01
130	5	26.3	335	14	Q83717	MATRIX PROTEIN.	6.00e+01	203	5	26.3	403	13	Q06726	STEROID RECEPTOR HOMOL	6.00e+01
131	5	26.3	335	14	Q83718	MATRIX PROTEIN.	6.00e+01	204	5	26.3	405	3	Q07413	HYPOETHICAL 46.7 KD P	6.00e+01
132	5	26.3	335	14	Q83721	MATRIX PROTEIN.	6.00e+01	205	5	26.3	409	2	Q54763	HYPOETHICAL 43.5 KD P	6.00e+01
133	5	26.3	335	14	Q83723	MATRIX PROTEIN.	6.00e+01	206	5	26.3	412	5	Q17743	COG1.1 PROTEIN.	6.00e+01
134	5	26.3	335	14	Q83535	MATRIX PROTEIN.	6.00e+01	207	5	26.3	412	5	Q18427	DNABJ.	6.00e+01
135	5	26.3	335	14	Q83538	MATRIX PROTEIN.	6.00e+01	208	5	26.3	413	14	Q11542	ENVELOPE GLYCOPROTEIN	6.00e+01
136	5	26.3	335	14	Q83529	MATRIX PROTEIN.	6.00e+01	209	5	26.3	413	2	Q33609	EPR PROTEIN.	6.00e+01
137	5	26.3	336	14	Q66410	STRUCTURAL PROTEIN.	6.00e+01	210	5	26.3	422	5	Q23846	PROTEIN TYROSINE KINAS	6.00e+01
138	5	26.3	336	10	Q96330	FLAVONOL SYNTHASE.	6.00e+01	211	5	26.3	426	5	Q22950	COSMID F08F3.	6.00e+01
139	5	26.3	338	2	Q47649	AGMATINE UREOHYDROLASE	6.00e+01	212	5	26.3	428	2	Q25459	HYPOETHICAL 47.9 KD P	6.00e+01
140	5	26.3	338	2	Q17167	ORF1 (BMC7).	6.00e+01	213	5	26.3	429	5	Q21523	M05D6.4 PROTEIN.	6.00e+01
141	5	26.3	339	2	Q66597	HISTIDINE KINASE SENSO	6.00e+01	214	5	26.3	434	13	Q93524	HYPOETHICAL 49.1 KD P	6.00e+01
142	5	26.3	339	11	Q88465	ZINC FINGER PROTEIN 10	6.00e+01	215	5	26.3	440	2	Q51728	CONSERVED HYPOTHETICAL	6.00e+01
143	5	26.3	340	5	Q21365	K08F8.5A PROTEIN.	6.00e+01	216	5	26.3	440	5	Q44637	F16B4.11 PROTEIN.	6.00e+01
144	5	26.3	341	6	Q02728	PREGNANCY-ASSOCIATED G	6.00e+01	217	5	26.3	444	5	Q18993	D2085.6 PROTEIN.	6.00e+01
145	5	26.3	341	6	Q46495	PREGNANCY-ASSOCIATED G	6.00e+01	218	5	26.3	444	2	Q67580	BIOTIN CARBOXYLASE.	6.00e+01
146	5	26.3	346	2	P71106	SIGB.	6.00e+01	219	5	26.3	465	5	Q18082	SIMILAR TO E. COLI ACY	6.00e+01
147	5	26.3	346	2	Q87364	HYPOTHETICAL 38.8 KD P	6.00e+01	220	5	26.3	467	6	P79802	PRESENTIN I.	6.00e+01
148	5	26.3	346	4	Q75477	KEO4P.	6.00e+01	221	5	26.3	469	3	Q06920	DNA FOR ORFS AND PPR1	6.00e+01
149	5	26.3	349	2	Q67215	HYPOTHETICAL 41.0 KD P	6.00e+01	222	5	26.3	472	10	Q22559	SIMILARITY TO PEPTIDAS	6.00e+01
150	5	26.3	352	2	Q47071	BEPE.	6.00e+01	223	5	26.3	472	2	Q87214	TRSF PROTEIN (TRAF).	6.00e+01
151	5	26.3	354	1	Q27328	SENSORY TRANSDUCTION H	6.00e+01	224	5	26.3	475	4	Q75255	PTPSIGA. PARTIAL CDS	6.00e+01
152	5	26.3	356	2	Q54910	OLIGOPEPTIDEPHEASE.	6.00e+01	225	5	26.3	475	10	Q23284	HYPOETHICAL 53.4 KD P	6.00e+01
153	5	26.3	357	10	Q42927	NARINGENIN 3-DIOXYGENA	6.00e+01	226	5	26.3	478	5	Q02168	SIMILAR TO AMP-ACTIVAT	6.00e+01
154	5	26.3	357	10	Q40344	FLAVONONE-3-HYDROXYLAS	6.00e+01	227	5	26.3	486	2	Q06491	GLU-TRNAGLIN AMIDOTRANS	6.00e+01
155	5	26.3	358	10	Q38877	FLAVONONE 3-HYDROXYLAS	6.00e+01	228	5	26.3	487	11	Q61518	KERATIN TYPE II (ENDO	6.00e+01
156	5	26.3	359	11	Q88819	ALPHA1.3-FUCOSYLTRANSFER	6.00e+01	229	5	26.3	490	11	Q61463	CYTOKERATIN ENDO A.	6.00e+01
157	5	26.3	360	3	P87045	OXYSTEROL-BINDING PROCT	6.00e+01	230	5	26.3	495	5	Q94224	SIMILARITY TO IG-LIKE	6.00e+01
158	5	26.3	361	10	Q04705	GIBBERELLIN 20-OXIDASE	6.00e+01	231	5	26.3	496	5	Q45245	H19N07.4 PROTEIN.	6.00e+01
159	5	26.3	363	2	P72668	HYPOTHETICAL 41.4 KD P	6.00e+01	232	5	26.3	498	10	Q39031	PROTEIN KINASE.	6.00e+01
160	5	26.3	363	5	Q76806	NATURATION INHIBITED P	6.00e+01	233	5	26.3	498	2	Q50179	ACYLTRANSFERASE (PUTAT	6.00e+01
161	5	26.3	365	10	Q04707	GIBBERELLIN 20-OXIDASE	6.00e+01	234	5	26.3	504	5	Q77432	EG:34F3.1 PROTEIN.	6.00e+01
162	5	26.3	365	10	Q04706	GIBBERELLIN 20-OXIDASE	6.00e+01	235	5	26.3	508	5	Q17099	AVL3-1.	6.00e+01
163	5	26.3	366	10	Q22530	FLAVANONE 3BETA-HYDROX	6.00e+01	236	5	26.3	509	2	P72957	HYPOETHICAL 58.5 KD P	6.00e+01
164	5	26.3	366	5	Q24830	CPN60 (FRAGMENT).	6.00e+01	237	5	26.3	510	2	Q66266	NADH OXIDASE/ALKYL HYD	6.00e+01
165	5	26.3	366	2	Q45378	DNA FOR LIPOPOLYSACCHA	6.00e+01	238	5	26.3	510	2	Q54469	NADH OXIDASE.	6.00e+01
166	5	26.3	366	10	Q23825	2-OXOGULUTARATE 3-DIOX	6.00e+01	239	5	26.3	515	13	Q42095	CGABP260 (FRAGMENT).	6.00e+01

240	5	26.3	534	2	045778	BFPD.	6.00e+01	313	5	26.3	837	5	076751	PUTATIVE ZINC METALLOP	6.00e+01
241	5	26.3	534	2	047070	BFPD.	6.00e+01	314	5	26.3	845	5	001914	F23H11.2 PROTEIN.	6.00e+01
242	5	26.3	536	5	055782	CHAPERONIN 60.	6.00e+01	315	5	26.3	846	13	037577	NCAM-140.	6.00e+01
243	5	26.3	539	10	080914	MYOSIN HEAVY CHAIN-LIK	6.00e+01	316	5	26.3	853	2	066463	HYPOTHETICAL 100.9 KD	6.00e+01
244	5	26.3	546	11	060457	ACAT.	6.00e+01	317	5	26.3	857	10	039837	LIPXYGENASE-3.	6.00e+01
245	5	26.3	549	2	025619	CONSERVED HYPOTHETICAL	6.00e+01	318	5	26.3	863	1	029837	SIGNAL-TRANSUCING HIS	6.00e+01
246	5	26.3	550	6	077760	STEROL O-ACYLTRANSFER	6.00e+01	319	5	26.3	875	5	025285	PHOSPHATIDYLINOSITOL-S	6.00e+01
247	5	26.3	550	6	077761	STEROL O-ACYLTRANSFER	6.00e+01	320	5	26.3	882	11	063029	ANDROGEN BINDING PROTE	6.00e+01
248	5	26.3	559	2	049563	DNA, TRANSPOSON-LIKE E	6.00e+01	321	5	26.3	891	2	035445	METHYL-ACCEPTING CHEMO	6.00e+01
249	5	26.3	562	4	033226	BS69 PROTEIN.	6.00e+01	322	5	26.3	904	13	031977	NMDA GLUTAMATE RECEPT	6.00e+01
250	5	26.3	562	5	023220	HMO82.3 PROTEIN.	6.00e+01	323	5	26.3	905	4	016253	PANCREATIC PEPTIDYLGLY	6.00e+01
251	5	26.3	565	5	017115	MEROZOITE SURFACE PROT	6.00e+01	324	5	26.3	913	14	039793	TYPE 1 OF IIIA GENE; S	6.00e+01
252	5	26.3	568	10	023174	HYPOTHETICAL 65.9 KD P	6.00e+01	325	5	26.3	917	4	015058	KIAA0350 (FRAGMENT).	6.00e+01
253	5	26.3	574	10	004545	SIMILAR TO ARABIDOPSIS	6.00e+01	326	5	26.3	922	11	082646	N-METHYL-D-ASPARTATE R	6.00e+01
254	5	26.3	581	2	051141	PENICILLIN-BINDING PRO	6.00e+01	327	5	26.3	935	13	031697	PEPTIDYLHYDROXYGLYCINE	6.00e+01
255	5	26.3	581	2	051317	PENICILLIN-BINDING PRO	6.00e+01	328	5	26.3	938	11	062683	NMDAR1 GLUTAMATE RECEP	6.00e+01
256	5	26.3	581	2	051107	PENICILLIN-BINDING PRO	6.00e+01	329	5	26.3	943	11	062648	N-METHYL-D-ASPARTATE R	6.00e+01
257	5	26.3	581	2	051099	PENICILLIN-BINDING PRO	6.00e+01	330	5	26.3	944	5	017411	HYPOTHETICAL 107.3 KD	6.00e+01
258	5	26.3	582	2	050960	PENICILLIN-BINDING PRO	6.00e+01	331	5	26.3	949	5	017399	HYPOTHETICAL 107.0 KD	6.00e+01
259	5	26.3	582	2	051120	PENICILLIN-BINDING PRO	6.00e+01	332	5	26.3	950	3	040452	D9461.10P.	6.00e+01
260	5	26.3	582	2	051097	PENICILLIN-BINDING PRO	6.00e+01	333	5	26.3	965	13	031979	N-METHYL-D-ASPARTATE R	6.00e+01
261	5	26.3	582	10	041099	NITRITE REDUCTASE PREC	6.00e+01	334	5	26.3	966	13	033338	N-METHYL-D-ASPARTATE R	6.00e+01
262	5	26.3	582	2	051233	PENICILLIN-BINDING PRO	6.00e+01	335	5	26.3	971	4	016252	PANCREATIC PEPTIDYLGLY	6.00e+01
263	5	26.3	582	2	051101	PENICILLIN-BINDING PRO	6.00e+01	336	5	26.3	971	2	051579	EXODEOXYRIBONUCLEASE V	6.00e+01
264	5	26.3	582	2	051098	PENICILLIN-BINDING PRO	6.00e+01	337	5	26.3	988	3	074681	CLASS II CHITIN SYNTHA	6.00e+01
265	5	26.3	584	2	051232	PENICILLIN-BINDING PRO	6.00e+01	338	5	26.3	1001	2	045344	TYPE III RESTRICTION E	6.00e+01
266	5	26.3	584	11	070573	TCF-3 PROTEIN.	6.00e+01	339	5	26.3	1003	11	070157	TOPOISOMERASE 3A (EC 5	6.00e+01
267	5	26.3	586	10	039183	SERINE/THREONINE PROTE	6.00e+01	340	5	26.3	1020	6	028350	PRECURSOR PEPTIDE.	6.00e+01
268	5	26.3	591	4	012867	N-METHYL-D-ASPARTATE R	6.00e+01	341	5	26.3	1047	10	037882	SUCROSE-PHOSPHATE SYNT	6.00e+01
269	5	26.3	600	2	074569	ASPARTATE KINASE.	6.00e+01	342	5	26.3	1051	2	070745	GLUCODEXTRANASE PRECUR	6.00e+01
270	5	26.3	605	2	083505	DNA PRIMASE (DNAP).	6.00e+01	343	5	26.3	1051	13	031048	KLK PRECURSOR.	6.00e+01
271	5	26.3	611	2	032975	NIFS-LIKE PROTEIN.	6.00e+01	344	5	26.3	1051	3	006680	CHROMOSOME IV COSMID 9	6.00e+01
272	5	26.3	617	14	083295	HEMAGGLUTININ.	6.00e+01	345	5	26.3	1100	13	037576	NCAM-180.	6.00e+01
273	5	26.3	633	5	022468	T13H5.3 PROTEIN.	6.00e+01	346	5	26.3	1136	1	058384	LTR-RETROTRANSPOSON SK	6.00e+01
274	5	26.3	635	5	018199	Y48E1B.5 PROTEIN.	6.00e+01	347	5	26.3	1136	1	058384	L136AA LONG HYPOTHETIC	6.00e+01
275	5	26.3	635	2	049936	PUTATIVE POLYKETIDE SY	6.00e+01	348	5	26.3	1139	1	059024	DNA POLYMERASE II LARG	6.00e+01
276	5	26.3	635	2	045820	CSPB GENE TRANSCRIPTION	6.00e+01	349	5	26.3	1155	1	038048	1155AA LONG HYPOTHETIC	6.00e+01
277	5	26.3	638	3	047910	PUTATIVE TRANSCRIPTION	6.00e+01	350	5	26.3	1261	11	084621	RECEPTOR-LINKED PROTEI	6.00e+01
278	5	26.3	643	3	012062	VERGICOLORIN B SYNTHAS	6.00e+01	351	5	26.3	1338	5	023927	PROTEIN TYROSINE KINAS	6.00e+01
279	5	26.3	647	3	059855	HEAT SHOCK PROTEIN.	6.00e+01	352	5	26.3	1360	14	055253	SPIKE GLYCOPROTEIN.	6.00e+01
280	5	26.3	649	10	081620	F8M12.9 PROTEIN.	6.00e+01	353	5	26.3	1361	14	083331	S GLYCOPROTEIN.	6.00e+01
281	5	26.3	650	2	048030	PUTATIVE ACCESSORY PRO	6.00e+01	354	5	26.3	1367	5	077354	MAL3P4.19 PROTEIN.	6.00e+01
282	5	26.3	650	2	048033	PUTATIVE ACCESSORY PRO	6.00e+01	355	5	26.3	1432	3	006585	HYPOTHETICAL 163.7 KD	6.00e+01
283	5	26.3	655	5	022196	SIMILARITY TO CCAAT DI	6.00e+01	356	5	26.3	1432	2	049932	PUTATIVE POLYKETIDE SY	6.00e+01
284	5	26.3	655	4	060620	KATANIN P80 SUBUNIT.	6.00e+01	357	5	26.3	1451	10	049142	POLYPROTEIN.	6.00e+01
285	5	26.3	660	5	094485	ORF DG1016 (FRAGMENT).	6.00e+01	358	5	26.3	1451	10	049140	POLYPROTEIN.	6.00e+01
286	5	26.3	660	2	059218	ARABINOSIDASE (EC 3.2.	6.00e+01	359	5	26.3	1451	10	049143	POLYPROTEIN.	6.00e+01
287	5	26.3	680	3	059765	HYPOTHETICAL 78.8 KD P	6.00e+01	360	5	26.3	1456	5	019127	SIMILAR TO C. ELEGANS	6.00e+01
288	5	26.3	696	2	067012	(P)PPGPP 3-PYRROPHOSPHO	6.00e+01	361	5	26.3	1461	13	012978	ENDODERMIN.	6.00e+01
289	5	26.3	703	2	005122	HAEMAGGLUTININ.	6.00e+01	362	5	26.3	1495	11	037808	PROTEIN-TYROSINE PHOSP	6.00e+01
290	5	26.3	703	3	008844	ORF YOR365C.	6.00e+01	363	5	26.3	1499	13	030815	PROTEIN-TYROSINE PHOSP	6.00e+01
291	5	26.3	704	5	061106	HYPOTHETICAL 80.4 KD P	6.00e+01	364	5	26.3	1501	11	064675	LEUKOCYTE COMMON ANTIG	6.00e+01
292	5	26.3	706	5	021666	SIMILARITY TO DROSOPHI	6.00e+01	365	5	26.3	1571	4	060469	DOWN SYNDROME CELL ADH	6.00e+01
293	5	26.3	710	5	093563	F2507.5 PROTEIN.	6.00e+01	366	5	26.3	1694	11	062230	SIALOADHESIN PRECURSOR	6.00e+01
294	5	26.3	713	2	086154	WLAF PROTEIN.	6.00e+01	367	5	26.3	1740	5	073734	SODIUM CHANNEL.	6.00e+01
295	5	26.3	717	14	040642	GLYCOPROTEIN H.	6.00e+01	368	5	26.3	1863	11	084605	LEUKOCYTE COMMON ANTIG	6.00e+01
296	5	26.3	739	6	028260	VASCULAR CELL ADHESION	6.00e+01	369	5	26.3	1888	11	088466	ZINC FINGER PROTEIN 10	6.00e+01
297	5	26.3	742	4	075156	KIAA0668 PROTEIN (FRAG	6.00e+01	370	5	26.3	1896	4	060468	DOWN SYNDROME CELL ADH	6.00e+01
298	5	26.3	742	2	050582	ALPHA-AMYLASE PRECURSO	6.00e+01	371	5	26.3	1898	11	064604	LEUKOCYTE COMMON ANTIG	6.00e+01
299	5	26.3	743	5	023869	D2 ORF.	6.00e+01	372	5	26.3	1904	11	064699	PROTEIN TYROSINE PHOSP	6.00e+01
300	5	26.3	745	5	021706	R04B5.9 PROTEIN.	6.00e+01	373	5	26.3	1907	11	064503	PROTEIN TYROSINE PHOSP	6.00e+01
301	5	26.3	759	2	025761	HYPOTHETICAL 88.8 KD P	6.00e+01	374	5	26.3	1910	5	022426	PROLINE-RICH.	6.00e+01
302	5	26.3	759	3	087244	HYPOTHETICAL 87.2 KD P	6.00e+01	375	5	26.3	1911	4	015718	PTPSIGNA PRECURSOR (EC	6.00e+01
303	5	26.3	768	3	042729	CELLOBIOSE DEHYDROGENA	6.00e+01	376	5	26.3	1948	4	013332	PROTEIN TYROSINE PHOSP	6.00e+01
304	5	26.3	775	10	048893	RESISTANCE PROTEIN CAN	6.00e+01	377	5	26.3	2030	5	061850	F55F10.2 PROTEIN.	6.00e+01
305	5	26.3	778	2	067326	OUTER MEMBRANE PROTEIN	6.00e+01	378	5	26.3	2030	8	032836	HYPOTHETICAL 245.6 KD	6.00e+01
306	5	26.3	779	5	019491	F16D3.2 PROTEIN.	6.00e+01	379	5	26.3	2207	11	062770	MUNC13-3.	6.00e+01
307	5	26.3	781	10	080661	T14N5.13 PROTEIN.	6.00e+01	380	5	26.3	2226	14	006502	POLYPROTEIN.	6.00e+01
308	5	26.3	793	4	044513	TYROSINE PHOSPHATASE P	6.00e+01	381	5	26.3	2327	5	016302	T05C3.2 PROTEIN.	6.00e+01
309	5	26.3	805	3	074239	RRM3/PIF1 HELICASE HOM	6.00e+01	382	5	26.3	2510	5	094658	EXPORTED SERINE/THREON	6.00e+01
310	5	26.3	819	5	090882	CODED FOR BY C. ELEGAN	6.00e+01	383	5	26.3	2553	5	022860	TALIN.	6.00e+01
311	5	26.3	822	13	031556	PROTEIN TYROSINE PHOSP	6.00e+01	384	5	26.3	3056	14	065892	POLYPROTEIN.	6.00e+01
312	5	26.3	828	5	000884	MIPAP (FRAGMENT).	6.00e+01	385	5	26.3	3056	14	065852	POLYPROTEIN.	6.00e+01

386	5	26.3	3212	5	094010	T08G11.1 PROTEIN.	6.00e+01	4	21.1	195	14	092484	P24-ACMPVP ORF129.	2.02e+03
387	5	26.3	3429	11	008614	CYTOSKELETAL PROTEIN.	6.00e+01	4	21.1	202	14	092164	ENVELOPE GLYCOPROTEIN	2.02e+03
388	4	21.1	26	14	091899	REV PROTEIN (FRAGMENT)	2.02e+03	4	21.1	204	14	092164	INTERLEUKIN-6 HOMOLOG	2.02e+03
389	4	21.1	36	8	092756	NADH DEHYDROGENASE SUB	2.02e+03	4	21.1	202	14	098823	TITIN (FRAGMENT).	2.02e+03
390	4	21.1	43	2	092733	EXOB (FRAGMENT).	2.02e+03	4	21.1	207	11	063580	TYPE 3 ACTIN (FRAGMENT	2.02e+03
391	4	21.1	56	14	092473	ACMPVP ORF116.	2.02e+03	4	21.1	210	10	041206	VERION PROTEIN.	2.02e+03
392	4	21.1	61	2	030739	THIOREDOXIN REDUCTASE	2.02e+03	4	21.1	210	14	089243	CONSERVED HYPOTHETICAL	2.02e+03
393	4	21.1	66	10	049211	REVERSE TRANSCRIPTASE	2.02e+03	4	21.1	212	5	029370	COSMID T22B7.	2.02e+03
394	4	21.1	66	2	096046	HYPOTHETICAL 7.7 KD PR	2.02e+03	4	21.1	212	1	023041	COSMID T22B7.	2.02e+03
395	4	21.1	68	7	019347	MHC CLASS II DO-ALPHA	2.02e+03	4	21.1	213	1	028410	HYPOTHETICAL 23.8 KD P	2.02e+03
396	4	21.1	73	2	068601	HYPOTHETICAL 8.0 KD PR	2.02e+03	4	21.1	214	13	042580	ER LUMEN PROTEIN RETAI	2.02e+03
397	4	21.1	83	14	092425	ACMPVP ORF60.	2.02e+03	4	21.1	217	10	036655	SUBUNIT OF PHOTOSYSTEM	2.02e+03
398	4	21.1	84	4	043262	LEUKEMIA ASSOCIATED GE	2.02e+03	4	21.1	219	14	073002	ENVELOPE GLYCOPROTEIN	2.02e+03
399	4	21.1	85	2	035030	YQQG PROTEIN.	2.02e+03	4	21.1	219	10	043536	ORF (FRAGMENT).	2.02e+03
400	4	21.1	89	14	092368	RNA POLYMERASE SUBUNIT	2.02e+03	4	21.1	222	14	012194	ENVELOPE GLYCOPROTEIN	2.02e+03
401	4	21.1	94	2	047771	ORF5.	2.02e+03	4	21.1	223	12	097967	70-KDA HEAT SHOCK PROT	2.02e+03
402	4	21.1	94	2	068670	GAS VESICLE PROTEIN GV	2.02e+03	4	21.1	225	11	064488	MPTDELTA PRECURSOR (E	2.02e+03
403	4	21.1	95	5	018687	F26E4.6 PROTEIN.	2.02e+03	4	21.1	225	14	068959	ORF 69.	2.02e+03
404	4	21.1	99	2	047101	REPRESSOR PROTEIN.	2.02e+03	4	21.1	226	7	030071	MHC CLASS II HLA-DQ-AL	2.02e+03
405	4	21.1	101	14	066904	PROTEIN 7A.	2.02e+03	4	21.1	228	14	096639	26KDA PROTEIN.	2.02e+03
406	4	21.1	103	2	071642	HYPOTHETICAL 11.7 KD P	2.02e+03	4	21.1	228	14	096638	26KDA PROTEIN.	2.02e+03
407	4	21.1	106	14	092397	TAT PROTEIN.	2.02e+03	4	21.1	228	14	065680	26KDA PROTEIN.	2.02e+03
408	4	21.1	107	14	012162	REV PROTEIN.	2.02e+03	4	21.1	228	14	065681	26KDA PROTEIN.	2.02e+03
409	4	21.1	108	3	012026	ORF YLR053C.	2.02e+03	4	21.1	229	2	070001	POTATIVE TWO-COMPONENT	2.02e+03
410	4	21.1	109	1	095916	ORF C01018.	2.02e+03	4	21.1	230	14	041133	AGS1L PROTEIN.	2.02e+03
411	4	21.1	110	14	092370	HYPOTHETICAL 12.7 KD P	2.02e+03	4	21.1	232	14	055495	P26.	2.02e+03
412	4	21.1	110	14	095682	CAPSID PROTEIN.	2.02e+03	4	21.1	232	5	022473	COSMID T14A8.	2.02e+03
413	4	21.1	114	13	079900	ACUTE PHASE SERUM AMYL	2.02e+03	4	21.1	233	14	090305	ENVELOPE PROTEIN.	2.02e+03
414	4	21.1	116	4	013138	MRNA CLONE WITH SIMILA	2.02e+03	4	21.1	233	1	026624	HYPOTHETICAL 26.5 KD P	2.02e+03
415	4	21.1	118	14	090488	ENVELOPE GLYCOPROTEIN	2.02e+03	4	21.1	235	2	06121	FLAGELLAR MOTOR PROTEI	2.02e+03
416	4	21.1	118	5	019925	COSMID F31A9.	2.02e+03	4	21.1	237	2	056003	SRRB.	2.02e+03
417	4	21.1	119	3	013530	YLR198CP.	2.02e+03	4	21.1	240	5	026368	IGLOO GROWTH ASSOCIATE	2.02e+03
418	4	21.1	120	14	086629	POLYPROTEIN PRECURSOR	2.02e+03	4	21.1	241	5	024968	CYST WALL PROTEIN 1 PR	2.02e+03
419	4	21.1	123	5	021019	F59A2.5 PROTEIN.	2.02e+03	4	21.1	241	2	025438	2-HYDROXY-6-OXOHEPTA-2	2.02e+03
420	4	21.1	128	4	099878	HISTONE H2A.	2.02e+03	4	21.1	242	2	047003	TRAK.	2.02e+03
421	4	21.1	129	14	092393	ACMPVP ORF26.	2.02e+03	4	21.1	242	2	046999	TRAK.	2.02e+03
422	4	21.1	134	14	057062	TAT PROTEIN.	2.02e+03	4	21.1	244	2	068988	CHLOROSOME ENVELOPE PR	2.02e+03
423	4	21.1	134	5	017495	ANTENNAL BINDING PROTE	2.02e+03	4	21.1	246	5	045242	CO4H5.4 PROTEIN.	2.02e+03
424	4	21.1	136	13	091825	EPIDERMAL PROTEIN (XEP	2.02e+03	4	21.1	248	14	089551	ORF20.	2.02e+03
425	4	21.1	141	3	005690	E-141 PROTEIN.	2.02e+03	4	21.1	250	14	036370	CAG PATHOGENICITY ISLA	2.02e+03
426	4	21.1	142	11	062358	MICROTUBULE-ASSOCIATED	2.02e+03	4	21.1	252	2	052625	SIMILARITY TO THE RGG	2.02e+03
427	4	21.1	142	3	042884	HYPOTHETICAL 15.9 KD P	2.02e+03	4	21.1	252	2	054977	29 KDA PROTEIN.	2.02e+03
428	4	21.1	142	14	041078	A596R PROTEIN.	2.02e+03	4	21.1	254	14	091256	ABC TRANSPORTER, ATP-B	2.02e+03
429	4	21.1	144	1	095921	ORF C01023.	2.02e+03	4	21.1	254	1	030306	MHC CLASS II HLA-DOAL.	2.02e+03
430	4	21.1	144	10	040437	RGP-3 (FRAGMENT).	2.02e+03	4	21.1	255	7	029882	HYPOTHETICAL PROTEIN M	2.02e+03
431	4	21.1	146	10	042178	SELENIUM BINDING PROTE	2.02e+03	4	21.1	258	1	058833	ORF-6.	2.02e+03
432	4	21.1	149	13	057313	PHOSPHOLIPASE A2 PRECU	2.02e+03	4	21.1	258	2	04780	41KBP FRAGMENT FROM LE	2.02e+03
433	4	21.1	155	10	022368	NACL-INDUCIBLE CA2+-BI	2.02e+03	4	21.1	259	14	087616	DNA ADENINE METHYLTRAN	2.02e+03
434	4	21.1	160	2	032662	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	261	7	030262	MHC CLASS II DM BETA-C	2.02e+03
435	4	21.1	160	2	032663	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	261	2	050274	POTATIVE LIPOPROTEIN (	2.02e+03
436	4	21.1	160	2	032668	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	261	2	050274	N-ACETYLURAMOVYL-L-ALA	2.02e+03
437	4	21.1	160	2	032673	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	263	14	089924	TRYPSIN-LIKE PROTEINASE.	2.02e+03
438	4	21.1	160	2	032679	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	265	3	074696	266AA LONG HYPOTHETICA	2.02e+03
439	4	21.1	160	2	032658	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	266	1	058108	LYSOPHOSPHOLIPASE.	2.02e+03
440	4	21.1	160	2	032656	VACUOLATING CYTOTOXIN	2.02e+03	4	21.1	266	1	028521	HYDROXYETHYLTHIAZOLE K	2.02e+03
441	4	21.1	161	14	089334	LECTIN HOMOLOG.	2.02e+03	4	21.1	268	1	082804	NUCLEIC ACID BINDING P	2.02e+03
442	4	21.1	162	1	028979	CONSERVED HYPOTHETICAL	2.02e+03	4	21.1	271	10	049216	ABO HISTO-BLOOD GROUP	2.02e+03
443	4	21.1	164	1	030322	L308_F2_61.	2.02e+03	4	21.1	274	6	018825	YFKH PROTEIN.	2.02e+03
444	4	21.1	167	2	049918	ENVELOPE GLYCOPROTEIN	2.02e+03	4	21.1	275	2	034437	RECEPTOR-LIKE KINASE (	2.02e+03
445	4	21.1	177	14	036313	ENVELOPE GLYCOPROTEIN	2.02e+03	4	21.1	275	10	081113	RGF PROTEIN, PUTATIVE.	2.02e+03
446	4	21.1	178	14	089776	CYSTATHIONINE ALPHA-SY	2.02e+03	4	21.1	282	2	087251	HYPOTHETICAL 32.4 KD P	2.02e+03
447	4	21.1	178	3	012633	LATENCY-RELATED OPEN R	2.02e+03	4	21.1	283	14	042090	RIBONUCLEASE H1.	2.02e+03
448	4	21.1	181	14	069259	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	286	4	060930	RET FINGER PROTEIN-LIK	2.02e+03
449	4	21.1	183	2	069732	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	288	10	082049	MITOCHONDRIAL CARRIER	2.02e+03
450	4	21.1	185	14	089918	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	290	5	062233	COSMID C06E2.	2.02e+03
451	4	21.1	185	14	089917	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	291	5	017721	HYPOPHETICAL PROTEIN M	2.02e+03
452	4	21.1	185	14	089919	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	292	1	058966	GP37-FUSOLIN-ACMPVP OR	2.02e+03
453	4	21.1	185	14	089920	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	292	14	092429	F22013.20.	2.02e+03
454	4	21.1	185	14	089921	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	295	10	064669	UL18 PROTEIN.	2.02e+03
455	4	21.1	185	14	089922	NUCLEOCAPSID PROTEIN (	2.02e+03	4	21.1	295	14	004530	HEMOGLOBIN RECEPTOR.	2.02e+03
456	4	21.1	188	14	090366	U131 PROTEIN (FRAGMENT	2.02e+03	4	21.1	296	2	072073	UBIQUITINATING ENZYME	2.02e+03
457	4	21.1	191	2	054635	HEAT SHOCK PROTEIN 70	2.02e+03	4	21.1	299	1	060800		
458	4	21.1	193	14	056882		2.02e+03	4	21.1	299	11			

532	4	21.1	300 14	O55284	MOVEMENT PROTEIN.	2.02e+03	605	4	21.1	380 2	P72936	REGULATORY COMPONENTS	2.02e+03
533	4	21.1	300 14	O55282	MOVEMENT PROTEIN.	2.02e+03	606	4	21.1	380 11	O55033	SH2/SH3 ADAPTOR PROTEI	2.02e+03
534	4	21.1	300 3	O74439	MITOCHONDRIAL CARRIER	2.02e+03	607	4	21.1	380 5	Q35514	SIMILAR TO SODIUM-DEPE	2.02e+03
535	4	21.1	304 1	O28479	ENDOGALACTINASE (CELM).	2.02e+03	608	4	21.1	382 8	Q35781	ATP SYNTHASE A CHAIN (	2.02e+03
536	4	21.1	306 1	O59375	306AA LONG HYPOTHETICA	2.02e+03	609	4	21.1	384 13	Q91322	PEPSINOGEN PRECURSOR.	2.02e+03
537	4	21.1	308 1	O29621	HYPOTHETICAL 33.7 KD P	2.02e+03	610	4	21.1	385 2	O33735	STREPTODORNASE.	2.02e+03
538	4	21.1	309 2	O85774	PUTATIVE RHIZOPINE PER	2.02e+03	611	4	21.1	385 2	P94599	SERINE/THREONINE KINAS	2.02e+03
539	4	21.1	309 2	O82849	RNA POLYMERASE SIGMA F	2.02e+03	612	4	21.1	387 3	O13918	PROBABLE ZINC/CADMIUM	2.02e+03
540	4	21.1	310 2	P96823	HYPOTHETICAL 34.0 KD P	2.02e+03	613	4	21.1	387 2	Q51395	MEMBRANE FUSION PROTEI	2.02e+03
541	4	21.1	311 2	P97031	HYPOTHETICAL 35.8 KD P	2.02e+03	614	4	21.1	388 2	P94288	ALKALINE D-PEPTIDASE.	2.02e+03
542	4	21.1	317 5	P91020	SIMILAR TO ALDEHYDE RE	2.02e+03	615	4	21.1	390 2	P74241	SULFATE ADENYLYLTRANSF	2.02e+03
543	4	21.1	319 4	O65407	POLYPROTEIN (FRAGMENTS	2.02e+03	616	4	21.1	391 2	O84574	GEN. SECRETION PROTEIN	2.02e+03
544	4	21.1	319 5	O17256	KO2E7.4 PROTEIN.	2.02e+03	617	4	21.1	395 2	O87998	PUTATIVE AMINO-SUGAR B	2.02e+03
545	4	21.1	321 5	O15757	PROTEIN PHOSPHATASE TY	2.02e+03	618	4	21.1	396 14	O86995	POLYPROTEIN PRECURSOR	2.02e+03
546	4	21.1	326 5	O26701	EXPRESSION SITE-ASSOCI	2.02e+03	619	4	21.1	397 10	O81815	MONOOXYGENASE.	2.02e+03
547	4	21.1	327 10	O49709	HYPOTHETICAL 36.8 KD P	2.02e+03	620	4	21.1	398 10	O23180	PATATIN HOMOLOG.	2.02e+03
548	4	21.1	327 5	O15560	ACTIN (FRAGMENT).	2.02e+03	621	4	21.1	399 3	Q00781	PLNTH.	2.02e+03
549	4	21.1	329 5	O76611	TP07H3.6 PROTEIN.	2.02e+03	622	4	21.1	399 2	P71464	PLNTH.	2.02e+03
550	4	21.1	330 2	O33240	HYPOTHETICAL 38.1 KD P	2.02e+03	623	4	21.1	401 2	P73176	REGULATORY COMPONENTS	2.02e+03
551	4	21.1	331 2	P72217	PSTS.	2.02e+03	624	4	21.1	402 2	O88124	PUTATIVE AMINOTRANSFER	2.02e+03
552	4	21.1	332 10	O63485	HYPOTHETICAL 37.7 KD P	2.02e+03	625	4	21.1	404 14	O89301	NUCLEOCAPSID PROTEIN (	2.02e+03
553	4	21.1	334 14	O36257	TAX PROTEIN (FRAGMENT)	2.02e+03	626	4	21.1	406 14	Q83136	POLYPROTEIN PRECURSOR	2.02e+03
554	4	21.1	335 4	Q13291	SIGNALING LYMPHOCYTIC	2.02e+03	627	4	21.1	407 1	O29048	PEPTIDE CHAIN RELEASE	2.02e+03
555	4	21.1	335 4	O43363	HXAL HUMAN HOMEOBOX PR	2.02e+03	628	4	21.1	409 2	O53088	ARGININE DEIMINASE (EC	2.02e+03
556	4	21.1	336 10	P93638	ACTIN (FRAGMENT).	2.02e+03	629	4	21.1	411 2	O50955	ANTIGEN P35, PUTATIVE	2.02e+03
557	4	21.1	336 10	P93636	ACTIN (FRAGMENT).	2.02e+03	630	4	21.1	413 2	O85779	DEFECTIVE IN FRUITING	2.02e+03
558	4	21.1	336 5	O19962	F32D8.1 PROTEIN.	2.02e+03	631	4	21.1	417 8	Q34184	NADH DEHYDROGENASE SUB	2.02e+03
559	4	21.1	336 10	P93639	ACTIN (FRAGMENT).	2.02e+03	632	4	21.1	420 2	O87350	LLAKR21 METHYLASE ENZY	2.02e+03
560	4	21.1	337 14	O93214	VP 10.	2.02e+03	633	4	21.1	421 3	O43068	HYPOTHETICAL 47.1 KD P	2.02e+03
561	4	21.1	338 2	O66495	HYPOTHETICAL 38.8 KD P	2.02e+03	634	4	21.1	428 3	O60140	HYPOTHETICAL 49.1 KD P	2.02e+03
562	4	21.1	339 2	O56136	SPAT, SIPB, SIPC, SIPD	2.02e+03	635	4	21.1	428 5	O16244	C13D9.3 PROTEIN.	2.02e+03
563	4	21.1	341 14	O90760	ANK3.	2.02e+03	636	4	21.1	429 14	O82136	NUCLEOCAPSID PROTEIN S	2.02e+03
564	4	21.1	341 2	O84037	CATIONIC AMINO ACID TR	2.02e+03	637	4	21.1	432 1	Q58874	HYPOTHETICAL PROTEIN M	2.02e+03
565	4	21.1	342 2	O83923	TP33 PROTEIN.	2.02e+03	638	4	21.1	433 8	O30071	ATP SYNTHASE BETA CHAI	2.02e+03
566	4	21.1	343 2	O67116	HYPOTHETICAL 41.1 KD P	2.02e+03	639	4	21.1	434 10	O65310	POLLENSESS3.	2.02e+03
567	4	21.1	344 10	O65921	PUTATIVE PROLINE-RICH	2.02e+03	640	4	21.1	437 14	P87601	41KBP FRAGMENT FROM LE	2.02e+03
568	4	21.1	346 1	O48941	MUTAGENIC FORNYLMETHA	2.02e+03	641	4	21.1	438 14	O36352	ENVELOPE PROTEIN.	2.02e+03
569	4	21.1	348 2	Q56033	SIPD.	2.02e+03	642	4	21.1	440 14	O69366	HOMOLOGUE OF HSV-1 77K	2.02e+03
570	4	21.1	349 11	O64316	ACTIN, BETA, CYTOPLASM	2.02e+03	643	4	21.1	441 5	O17304	2K250.8 PROTEIN.	2.02e+03
571	4	21.1	350 11	O60686	MICROTUBULE-ASSOCIATED	2.02e+03	644	4	21.1	442 4	Q13703	ALPHA1-ANTICRYMOTRYP	2.02e+03
572	4	21.1	350 5	O25026	CYSTINE PROTEINASE.	2.02e+03	645	4	21.1	443 5	O45422	F31D4.7 PROTEIN.	2.02e+03
573	4	21.1	350 5	O26514	CALRETICULIN (FRAGMENT	2.02e+03	646	4	21.1	443 8	O33559	MURF1 PROTEIN (FRAGMEN	2.02e+03
574	4	21.1	351 3	O42645	PUTATIVE ALPHA-TUBULIN	2.02e+03	647	4	21.1	444 14	O89335	(FRAGMENT).	2.02e+03
575	4	21.1	352 2	O66715	HETERODISULFIDE REDUCT	2.02e+03	648	4	21.1	445 10	O81998	TRANSCRIPTION FACTOR V	2.02e+03
576	4	21.1	353 14	O92269	NIB PROTEIN/COAT PROTE	2.02e+03	649	4	21.1	445 14	Q86104	SU PROTEIN (FRAGMENT).	2.02e+03
577	4	21.1	354 14	O65168	PB354L.	2.02e+03	650	4	21.1	445 14	O86105	SU PROTEIN (FRAGMENT).	2.02e+03
578	4	21.1	355 14	O67638	GLYCOPROTEIN HOMOLOG I	2.02e+03	651	4	21.1	445 14	O86103	SU PROTEIN (FRAGMENT).	2.02e+03
579	4	21.1	355 14	O88276	COAT PROTEIN (FRAGMENT	2.02e+03	652	4	21.1	445 5	O19730	F22D6.11 PROTEIN.	2.02e+03
580	4	21.1	358 10	Q39103	GM4.	2.02e+03	653	4	21.1	446 1	O27452	HYPOTHETICAL 48.0 KD P	2.02e+03
581	4	21.1	359 2	O34032	ORF1.	2.02e+03	654	4	21.1	448 14	O39254	COUNTERPART OF HSV-1 G	2.02e+03
582	4	21.1	360 5	O26446	CYCLIC AMP-DEPENDENT P	2.02e+03	655	4	21.1	448 10	O24399	FIMBRIN/PLASTIN-LIKE (	2.02e+03
583	4	21.1	362 2	O52368	PERIPLASMIC SUGAR BIND	2.02e+03	656	4	21.1	450 3	Q08448	CHROMOSOME XV READING	2.02e+03
584	4	21.1	362 1	O27988	CYTOCHROME OXIDASE,SUB	2.02e+03	657	4	21.1	451 3	O43076	HYPOTHETICAL 52.4 KD P	2.02e+03
585	4	21.1	362 14	O92542	POLYPROTEIN (FRAGMENT)	2.02e+03	658	4	21.1	453 10	O49285	F22K20.9 PROTEIN.	2.02e+03
586	4	21.1	364 2	O67641	HYPOTHETICAL 40.8 KD P	2.02e+03	659	4	21.1	453 3	Q08601	CHROMOSOME XV READING	2.02e+03
587	4	21.1	364 10	O46461	F22O13.12.	2.02e+03	660	4	21.1	457 14	O91894	ENVELOPE GLYCOPROTEIN	2.02e+03
588	4	21.1	364 14	O90096	HYPOPHETICAL 40.6 KD P	2.02e+03	661	4	21.1	457 14	Q65872	ORF5 (FRAGMENT).	2.02e+03
589	4	21.1	365 10	O23238	PYRUVATE DEHYDROGENASE	2.02e+03	662	4	21.1	459 5	P91876	OVARIAN SPECIFIC SERIN	2.02e+03
590	4	21.1	366 10	O82657	POLYPROTEIN (FRAGMENT)	2.02e+03	663	4	21.1	461 14	O89329	POLYPROTEIN PRECURSOR	2.02e+03
591	4	21.1	367 14	O89736	POLYPROTEIN (FRAGMENT)	2.02e+03	664	4	21.1	461 2	O69539	HYPOTHETICAL 47.2 KD P	2.02e+03
592	4	21.1	367 14	O89735	POLYPROTEIN (FRAGMENT)	2.02e+03	665	4	21.1	462 2	O25455	BILARY GLYCOPROTEIN.	2.02e+03
593	4	21.1	368 5	O17742	EO3H4.8 PROTEIN.	2.02e+03	666	4	21.1	462 2	Q13854	CONSERVED HYPOTHETICAL	2.02e+03
594	4	21.1	371 1	O30865	XYLENE MONOOXYGENASE H	2.02e+03	667	4	21.1	464 10	Q40073	RIBULOSE 1,5-BISPHOSPH	2.02e+03
595	4	21.1	371 1	O59539	371AA LONG HYPOTHETICA	2.02e+03	668	4	21.1	465 10	O24359	GLUCOSE-6-PHOSPHATE 1-	2.02e+03
596	4	21.1	371 3	O13838	HYPOTHETICAL 40.3 KD P	2.02e+03	669	4	21.1	470 10	O82655	HYPOTHETICAL 53.9 KD P	2.02e+03
597	4	21.1	374 10	O24627	DEFECTIVE GIBBERELLIN	2.02e+03	670	4	21.1	472 10	Q43327	RIBULOSE BISPHOSPHATE	2.02e+03
598	4	21.1	374 11	O63677	MICROTUBULE ASSOCIATED	2.02e+03	671	4	21.1	473 5	Q23233	COSMID Y102E9.	2.02e+03
599	4	21.1	375 13	O73815	BETA-ACTIN.	2.02e+03	672	4	21.1	474 8	O47283	RIBULOSE BISPHOSPHATE	2.02e+03
600	4	21.1	376 3	O60104	HYPOTHETICAL 43.5 KD P	2.02e+03	673	4	21.1	475 2	O69455	P50 ADHESIN PRECURSOR.	2.02e+03
601	4	21.1	377 10	O65315	ACTIN.	2.02e+03	674	4	21.1	475 5	O21555	M18.3 PROTEIN.	2.02e+03
602	4	21.1	377 10	P93775	ACTIN.	2.02e+03	675	4	21.1	481 10	O64814	F21P24.11 PROTEIN.	2.02e+03
603	4	21.1	378 2	O60012	NAD ALCOHOL DEHYDROGEN	2.02e+03	676	4	21.1	482 14	O92516	FIRST TRIPLE GENE BLOC	2.02e+03
604	4	21.1	379 10	O40994	GIBBERELLIN 20-OXIDASE	2.02e+03	677	4	21.1	483 5	O77145	AKT/PKB SERINE/THREONI	2.02e+03

678	4	21.1	484	10	Q04218	TNP1 PROTEIN.	2.02e+03	751	4	21.1	623	11	Q62005	ZONA PELLUCIDA GLYCOP	2.02e+03
679	4	21.1	486	1	Q58792	HYPOTHETICAL PROTEIN M	2.02e+03	752	4	21.1	631	10	Q04510	SEQUENCE OF BAC F21M12	2.02e+03
680	4	21.1	488	5	P06624	MYOSIN (FRAGMENT).	2.02e+03	753	4	21.1	631	9	Q37967	TELOMERASE.	2.02e+03
681	4	21.1	489	10	Q65209	1-AMINOCYCLOPROPANE-1-	2.02e+03	754	4	21.1	632	5	Q20447	F46B6_6 PROTEIN.	2.02e+03
682	4	21.1	491	2	P96563	HYPOTHETICAL HYDROXYLA	2.02e+03	755	4	21.1	632	2	Q02840	BETA-N-ACETYLGLUCOSAMI	2.02e+03
683	4	21.1	493	14	Q93184	HYPOTHETICAL 54.5 KD P	2.02e+03	756	4	21.1	635	14	Q98262	MC095R.	2.02e+03
684	4	21.1	494	14	Q84389	CAPSID PROTEIN (FRAGME	2.02e+03	757	4	21.1	637	14	Q91320	VETF-1, EARLY TRANSCRI	2.02e+03
685	4	21.1	494	14	Q84390	CAPSID PROTEIN (FRAGME	2.02e+03	758	4	21.1	647	3	Q12045	CHROMOSOME XVI READING	2.02e+03
686	4	21.1	498	8	Q31931	ADENOSINE TRIPHOSPHATA	2.02e+03	759	4	21.1	648	2	Q34507	YLOP PROTEIN.	2.02e+03
687	4	21.1	499	5	Q44658	B0213_14 PROTEIN.	2.02e+03	760	4	21.1	651	5	Q26574	FIMBRIN.	2.02e+03
688	4	21.1	506	14	Q85448	PROTEIN 7.	2.02e+03	761	4	21.1	660	10	Q64380	POLYA BINDING PROTEIN	2.02e+03
689	4	21.1	507	14	Q57126	L1 PROTEIN.	2.02e+03	762	4	21.1	666	14	Q91259	OVERLAPPING PROTEIN.	2.02e+03
690	4	21.1	509	11	Q88548	TYPE II PNEUMOCYTE CD3	2.02e+03	763	4	21.1	668	14	P89514	NONSTRUCTURAL PROTEIN	2.02e+03
691	4	21.1	509	14	Q86566	NUCLEOPROTEIN (NP).	2.02e+03	764	4	21.1	668	14	P89515	NONSTRUCTURAL PROTEIN	2.02e+03
692	4	21.1	512	2	P94684	GLUCOSE-6-PHOSPHATE 1-	2.02e+03	765	4	21.1	668	14	P89512	NONSTRUCTURAL PROTEIN	2.02e+03
693	4	21.1	514	14	Q88443	COMPLETE GENOME.	2.02e+03	766	4	21.1	668	14	P89516	DNA, COMPLETE GENOME S	2.02e+03
694	4	21.1	519	14	Q57045	IMMEDIATE EARLY PROTEI	2.02e+03	767	4	21.1	668	14	P89516	NONSTRUCTURAL PROTEIN	2.02e+03
695	4	21.1	526	10	Q82783	NUS RECEPTOR.	2.02e+03	768	4	21.1	668	14	P89513	NONSTRUCTURAL PROTEIN	2.02e+03
696	4	21.1	527	2	Q25521	TYPE I RESTRICTION ENZ	2.02e+03	769	4	21.1	668	2	P77864	ENDOGUCANASE D PRECUR	2.02e+03
697	4	21.1	528	2	Q83650	LYSL-TRNA SYNTHETASE	2.02e+03	770	4	21.1	672	14	Q83429	NONSTRUCTURAL PROTEIN	2.02e+03
698	4	21.1	528	2	Q60158	RNA POLYMERASE SIGMA F	2.02e+03	771	4	21.1	673	14	Q92390	ACMNPV ORE23.	2.02e+03
699	4	21.1	528	2	Q60162	RNA POLYMERASE SIGMA F	2.02e+03	772	4	21.1	676	5	Q44188	F33D4_1 PROTEIN.	2.02e+03
700	4	21.1	529	2	Q50747	EFFLUX PROTEIN.	2.02e+03	773	4	21.1	680	14	Q98187	MC018U.	2.02e+03
701	4	21.1	530	14	Q90713	VP2 PROTEIN (FRAGMENT)	2.02e+03	774	4	21.1	688	14	Q65575	PUTATIVE VIRION PROTEI	2.02e+03
702	4	21.1	532	2	Q30727	CYTCHROME CB3 OXIDAS	2.02e+03	775	4	21.1	693	14	Q69090	ULAT7.	2.02e+03
703	4	21.1	535	14	Q36607	POLYPROTEIN (FRAGMENT)	2.02e+03	776	4	21.1	694	5	Q94916	DF22.	2.02e+03
704	4	21.1	535	14	Q90128	STRUCTURAL PROTEIN VP2	2.02e+03	777	4	21.1	706	2	Q67434	HYPOTHETICAL 83.1 KD P	2.02e+03
705	4	21.1	539	2	Q51679	CCON (EC 1.9.3.1).	2.02e+03	778	4	21.1	720	14	Q36332	ENVELOPE PROTEIN.	2.02e+03
706	4	21.1	542	14	Q11448	VP78.	2.02e+03	779	4	21.1	722	14	Q84394	DNA, COMPLETE GENOME S	2.02e+03
707	4	21.1	543	14	Q44130	F45E4_11 PROTEIN.	2.02e+03	780	4	21.1	723	11	Q55166	ARE1.	2.02e+03
708	4	21.1	544	1	Q29452	L-LACTATE PEROXASE (LC	2.02e+03	781	4	21.1	723	14	Q84207	CODING SEQUENCE.	2.02e+03
709	4	21.1	544	5	Q02081	SIMILARITY TO A SHORT	2.02e+03	782	4	21.1	733	13	Q12944	PUTATIVE RECOMBINATION	2.02e+03
710	4	21.1	545	14	Q90054	ORE2.	2.02e+03	783	4	21.1	735	2	P94180	ADENYLATE CYCLASE (EC	2.02e+03
711	4	21.1	547	5	Q18895	SIMILAR TO OTHER MEMBE	2.02e+03	784	4	21.1	737	5	Q17305	GOLD9.2 PROTEIN.	2.02e+03
712	4	21.1	550	14	Q82498	HENAGGLUTININS HAI AND	2.02e+03	785	4	21.1	744	10	Q50037	CHROMOMETHYLASE (FRAGM	2.02e+03
713	4	21.1	550	2	P95752	FIBRONECTIN-BINDING PR	2.02e+03	786	4	21.1	745	3	P79086	TWO-COMPONENT HISTIDIN	2.02e+03
714	4	21.1	553	2	Q11153	RIBOSOMAL PROTEIN S1 (	2.02e+03	787	4	21.1	745	2	Q86648	HYPOTHETICAL 81.0 KD P	2.02e+03
715	4	21.1	559	11	Q35533	CARBOXYLESTERASE PRECU	2.02e+03	788	4	21.1	745	10	P93041	HOMEODOMAIN PROTEIN AH	2.02e+03
716	4	21.1	560	14	Q06347	REVERSE TRANSCRIPTASE	2.02e+03	789	4	21.1	751	14	Q03725	COAT PROTEIN VP1 [CONT	2.02e+03
717	4	21.1	560	10	Q43319	3-HYDROXY-3-METHYLGLUT	2.02e+03	790	4	21.1	751	14	Q03726	COAT PROTEIN VP1 [CONT	2.02e+03
718	4	21.1	560	2	Q07372	LIPOPROTEIN.	2.02e+03	791	4	21.1	756	14	Q90155	RNA-DEPENDENT RNA POLY	2.02e+03
719	4	21.1	560	14	Q36433	NUCLEOPROTEIN.	2.02e+03	792	4	21.1	757	5	Q26126	MICRONEME PROTEIN-1 (F	2.02e+03
720	4	21.1	560	10	Q49141	CHROMOMETHYLASE.	2.02e+03	793	4	21.1	757	5	Q26127	MICRONEME PROTEIN-1 (F	2.02e+03
721	4	21.1	565	14	Q86639	HENAGGLUTININ.	2.02e+03	794	4	21.1	767	5	Q18103	COSMID C18H2.	2.02e+03
722	4	21.1	565	5	Q21615	RO1E6_4 PROTEIN.	2.02e+03	795	4	21.1	768	5	Q21523	RO1H10_7.	2.02e+03
723	4	21.1	567	5	Q16108	FORK HEAD 1.	2.02e+03	796	4	21.1	774	2	Q05258	HYPOTHETICAL 86.4 KD P	2.02e+03
724	4	21.1	568	14	Q84419	PBCV-1 HYALURONIC ACID	2.02e+03	797	4	21.1	776	14	Q86179	VP4.	2.02e+03
725	4	21.1	571	11	Q55160	IP63 PROTEIN.	2.02e+03	798	4	21.1	776	14	Q86191	VP4 PROTEIN.	2.02e+03
726	4	21.1	571	2	Q53893	ABSAL.	2.02e+03	799	4	21.1	776	14	Q98637	VP4.	2.02e+03
727	4	21.1	572	2	Q45469	BURKHOLDERIA SOLANACEA	2.02e+03	800	4	21.1	776	14	Q86180	VP4.	2.02e+03
728	4	21.1	574	14	P88811	FUSION GLYCOPROTEIN PR	2.02e+03	801	4	21.1	776	14	Q98636	VP4.	2.02e+03
729	4	21.1	574	14	Q90698	F PROTEIN.	2.02e+03	802	4	21.1	777	2	P72644	PHOSPHORIBOSYLFORMYL G	2.02e+03
730	4	21.1	574	14	Q09720	FUSION PROTEIN (F).	2.02e+03	803	4	21.1	777	5	Q94235	CODED FOR BY C. ELEGAN	2.02e+03
731	4	21.1	574	14	Q09635	FUSION PROTEIN (F).	2.02e+03	804	4	21.1	784	4	Q06053	TOLL/INTERLEUKIN-1 REC	2.02e+03
732	4	21.1	579	14	Q84396	CAPSID PROTEIN VP2.	2.02e+03	805	4	21.1	787	14	Q98805	HOMOLOG OF VACCINIA VI	2.02e+03
733	4	21.1	579	10	Q42947	DEHYDROQUINATE DEHYDRA	2.02e+03	806	4	21.1	787	2	Q06776	HYPOTHETICAL 86.1 KD P	2.02e+03
734	4	21.1	584	14	Q93187	CAPSID PROTEIN.	2.02e+03	807	4	21.1	789	14	Q06120	GENOME POLYPROTEIN [CO	2.02e+03
735	4	21.1	584	14	P87701	CAPSID PROTEIN (VP2).	2.02e+03	808	4	21.1	795	5	Q18394	NICOTINIC ACETYLCHOLIN	2.02e+03
736	4	21.1	584	14	P89517	CAPSID PROTEIN 2.	2.02e+03	809	4	21.1	797	14	Q03810	ENVELOPE POLYPROTEIN G	2.02e+03
737	4	21.1	584	14	P90438	CAPSID PROTEIN 2.	2.02e+03	810	4	21.1	804	2	Q84193	ENVELOPE POLYPROTEIN G	2.02e+03
738	4	21.1	584	14	P89947	CAPSID PROTEIN (VP2).	2.02e+03	811	4	21.1	810	14	Q90053	DNA GYRASE SUBUNIT B.	2.02e+03
739	4	21.1	584	14	Q96801	CAPSID PROTEIN VP2.	2.02e+03	812	4	21.1	824	4	Q13099	ORF1.	2.02e+03
740	4	21.1	584	14	P90445	CAPSID PROTEIN VP2.	2.02e+03	813	4	21.1	828	13	Q91743	TG737.	2.02e+03
741	4	21.1	589	2	P72364	HYPOTHETICAL 65.9 KD P	2.02e+03	814	4	21.1	829	10	Q64584	FIBROBLAST GROWTH FACT	2.02e+03
742	4	21.1	589	2	Q06988	HYPOTHETICAL 68.9 KD P	2.02e+03	815	4	21.1	830	10	Q64584	T29F13.6 PROTEIN.	2.02e+03
743	4	21.1	593	4	Q76083	CGMP PHOSPHODIESTERASE	2.02e+03	816	4	21.1	830	10	Q04864	1,4-ALPHA-GLUCAN BRANC	2.02e+03
744	4	21.1	598	2	Q88138	TPR PROTEIN D (TPRD).	2.02e+03	817	4	21.1	843	14	P87603	41KBP FRAGMENT FROM LE	2.02e+03
745	4	21.1	599	2	Q13406	REGULATORY PROTEIN (FR	2.02e+03	818	4	21.1	844	2	Q64804	ETHYLENE RESPONSE SENS	2.02e+03
746	4	21.1	607	10	Q39775	LEGUMIN, 11S GLOBULIN.	2.02e+03	819	4	21.1	845	14	Q36329	ETHYLENE RESPONSE SENS	2.02e+03
747	4	21.1	611	2	Q69514	DNA POLYMERASE III SUB	2.02e+03	820	4	21.1	845	14	Q36330	ENVELOPE PROTEIN.	2.02e+03
748	4	21.1	615	5	Q20845	SIMILAR TO SER/THR PRO	2.02e+03	821	4	21.1	845	14	Q36331	ENVELOPE PROTEIN.	2.02e+03
749	4	21.1	621	2	Q07653	CELLULASE B.	2.02e+03	822	4	21.1	845	14	Q42056	ENVELOPE PROTEIN.	2.02e+03
750	4	21.1	622	2	Q31248	PUTATIVE PEPTIDYL-PROL	2.02e+03	823	4	21.1	845	14	Q36326	ENVELOPE PROTEIN.	2.02e+03

824	4	21.1	845 14	036328	ENVELOPE PROTEIN.	2.02e+03	897	4	21.1	1178 4	075857	NEURONAL APOPTOSIS INH	2.02e+03
825	4	21.1	845 14	036327	ENVELOPE PROTEIN.	2.02e+03	898	4	21.1	1188 11	P97500	ZINC FINGER PROTEIN PN	2.02e+03
826	4	21.1	845 14	036319	ENVELOPE PROTEIN.	2.02e+03	899	4	21.1	1188 14	P90434	134K PROTEIN.	2.02e+03
827	4	21.1	845 14	036315	ENVELOPE PROTEIN.	2.02e+03	900	4	21.1	1191 3	042765	ELONGATION-LIKE FACTOR	2.02e+03
828	4	21.1	845 14	036318	ENVELOPE PROTEIN.	2.02e+03	901	4	21.1	1202 11	P70313	NITRIC OXIDE SYNTHASE	2.02e+03
829	4	21.1	847 2	055476	SOLUBLE LYTC TRANSGLY	2.02e+03	902	4	21.1	1203 5	077312	MAL3P1.9 PROTEIN.	2.02e+03
830	4	21.1	853 14	089469	ENVELOPE POLYPROTEIN.	2.02e+03	903	4	21.1	1210 13	092137	PHOSPHOLIPASE C BETA T	2.02e+03
831	4	21.1	853 10	043446	LIPOXYGENASE L-5.	2.02e+03	904	4	21.1	1217 5	017889	F54F12.1 PROTEIN (EC 3	2.02e+03
832	4	21.1	855 5	P90349	ATP-DEPENDENT RNA HELI	2.02e+03	905	4	21.1	1218 5	023326	SIMILAR TO MAP KINASE	2.02e+03
833	4	21.1	858 14	089473	ENVELOPE POLYPROTEIN.	2.02e+03	906	4	21.1	1247 5	017461	B0240.2 PROTEIN.	2.02e+03
834	4	21.1	859 14	036354	ENVELOPE PROTEIN.	2.02e+03	907	4	21.1	1254 3	074502	MUTS FAMILY DNA MISMAT	2.02e+03
835	4	21.1	859 14	036351	ENVELOPE PROTEIN.	2.02e+03	908	4	21.1	1270 14	093163	145 KDA PROTEIN.	2.02e+03
836	4	21.1	859 14	036350	ENVELOPE PROTEIN.	2.02e+03	909	4	21.1	1270 13	042287	INTERSECTIN.	2.02e+03
837	4	21.1	859 14	036334	ENVELOPE PROTEIN.	2.02e+03	910	4	21.1	1277 11	035821	PAR INTERACTING PROTEI	2.02e+03
838	4	21.1	859 14	036337	ENVELOPE PROTEIN.	2.02e+03	911	4	21.1	1280 5	022554	SIMILARITY TO S. CREVEV	2.02e+03
839	4	21.1	859 14	036338	ENVELOPE PROTEIN.	2.02e+03	912	4	21.1	1286 14	057204	RNA POLYMERASE SUBUNIT	2.02e+03
840	4	21.1	859 14	092814	ENVELOPE POLYPROTEIN.	2.02e+03	913	4	21.1	1289 14	085289	DNA-DEPENDENT RNA POLY	2.02e+03
841	4	21.1	859 14	036324	ENVELOPE PROTEIN.	2.02e+03	914	4	21.1	1289 14	098246	MC079R.	2.02e+03
842	4	21.1	859 14	036325	ENVELOPE PROTEIN.	2.02e+03	915	4	21.1	1318 14	066542	P140.	2.02e+03
843	4	21.1	859 14	036356	ENVELOPE PROTEIN.	2.02e+03	916	4	21.1	1333 11	070546	UBIQUITOUSLY TRANSCRIB	2.02e+03
844	4	21.1	859 14	036333	ENVELOPE PROTEIN.	2.02e+03	917	4	21.1	1336 5	094479	ABCA (FRAGMENT).	2.02e+03
845	4	21.1	859 14	036336	ENVELOPE PROTEIN.	2.02e+03	918	4	21.1	1344 14	098651	POLYPROTEIN (FRAGMENTS	2.02e+03
846	4	21.1	859 14	036335	ENVELOPE PROTEIN.	2.02e+03	919	4	21.1	1350 5	027571	NITRIC OXIDE SYNTHASE	2.02e+03
847	4	21.1	859 14	036322	ENVELOPE PROTEIN.	2.02e+03	920	4	21.1	1361 14	066199	SURFACE PROTEIN.	2.02e+03
848	4	21.1	859 14	036323	ENVELOPE PROTEIN.	2.02e+03	921	4	21.1	1368 5	025661	PUTATIVE GLUTAMATE AND	2.02e+03
849	4	21.1	859 14	036316	ENVELOPE PROTEIN.	2.02e+03	922	4	21.1	1377 14	083137	POLYPEPTIDE (FRAGMENT)	2.02e+03
850	4	21.1	860 14	036333	ENVELOPE PROTEIN.	2.02e+03	923	4	21.1	1387 5	020766	HYPOTHETICAL 160.4 KD	2.02e+03
851	4	21.1	868 3	007798	CHROMOSOME XII READING	2.02e+03	924	4	21.1	1403 2	044494	MANNURONAN C-5-EPIMERA	2.02e+03
852	4	21.1	873 14	089274	ZINC FINGER PROTEIN.	2.02e+03	925	4	21.1	1405 11	061202	NUCLEAR RECEPTOR COACT	2.02e+03
853	4	21.1	873 5	023717	ATP OPERON (FRAGMENT).	2.02e+03	926	4	21.1	1408 1	027557	CELL SURFACE GLYCOPROT	2.02e+03
854	4	21.1	885 2	053031	HERPESVIRUS TYPE 6 (HH	2.02e+03	927	4	21.1	1453 14	091603	SPIKE PROTEIN.	2.02e+03
855	4	21.1	893 14	069062	GAG PRO POL POLYPROTEI	2.02e+03	928	4	21.1	1465 5	017909	H06001.2 PROTEIN.	2.02e+03
856	4	21.1	899 14	083393	HEXON.	2.02e+03	929	4	21.1	1520 10	039160	MYOSIN.	2.02e+03
857	4	21.1	899 5	044872	R06B10.4 PROTEIN.	2.02e+03	930	4	21.1	1555 14	085274	POLYPROTEIN (FRAGMENT)	2.02e+03
858	4	21.1	903 14	040957	HEXON.	2.02e+03	931	4	21.1	1559 10	049889	MODIFICATION METHYLASE	2.02e+03
859	4	21.1	909 10	065027	HOMOSERINE DEHYDROGENA	2.02e+03	932	4	21.1	1583 14	090304	PEPOMER PROTEIN.	2.02e+03
860	4	21.1	912 4	052888	GUANINE NUCLEOTIDE EXC	2.02e+03	933	4	21.1	1642 5	062055	CO9F9.2 PROTEIN.	2.02e+03
861	4	21.1	919 14	067674	POLYPROTEIN PRECURSOR	2.02e+03	934	4	21.1	1654 2	053047	OUTER MEMBRANE PROTEIN	2.02e+03
862	4	21.1	919 10	024375	ALPHA-GLUCOSIDASE (EC	2.02e+03	935	4	21.1	1751 14	092815	PR GAG-PRO-POL.	2.02e+03
863	4	21.1	935 2	031000	INTMIN.	2.02e+03	936	4	21.1	1789 14	083883	NONSTRUCTURAL POLYPROT	2.02e+03
864	4	21.1	942 5	018298	C29E6.4 PROTEIN.	2.02e+03	937	4	21.1	1811 14	036184	NONSTRUCTURAL POLYPROT	2.02e+03
865	4	21.1	955 14	086782	L2-VP2 (SEGMENT 2).	2.02e+03	938	4	21.1	1897 14	085406	(STRAIN BANGLADESH-197	2.02e+03
866	4	21.1	961 5	009981	HYPOTHETICAL 107.4 KD	2.02e+03	939	4	21.1	1943 11	061307	ANKYRIN 3 (ANKYRIN G)	2.02e+03
867	4	21.1	962 14	065162	PC962R.	2.02e+03	940	4	21.1	1956 5	004956	ATPASE 1.	2.02e+03
868	4	21.1	963 3	074911	RAD16 NUCLEOTIDE EXCIS	2.02e+03	941	4	21.1	1957 4	015064	KIAA0357 PROTEIN (FRAG	2.02e+03
869	4	21.1	965 3	013992	HYPOTHETICAL 110.9 KD	2.02e+03	942	4	21.1	1965 5	061893	T07H8.4 PROTEIN.	2.02e+03
870	4	21.1	972 14	083862	106.4KD PROTEIN.	2.02e+03	943	4	21.1	1993 5	P90670	SODIUM CHANNEL ALPHA-S	2.02e+03
871	4	21.1	982 5	017583	C09H6.2 PROTEIN.	2.02e+03	944	4	21.1	2023 14	091632	RNA-DEPENDENT RNA POLY	2.02e+03
872	4	21.1	983 14	069532	HERPESVIRUS 6 IMMEDIAT	2.02e+03	945	4	21.1	2055 11	088938	RHO/RAC-INTERACTING CI	2.02e+03
873	4	21.1	1010 5	P90885	F54C8.3 PROTEIN.	2.02e+03	946	4	21.1	2109 14	086125	POLYMERASE (L PROTEIN)	2.02e+03
874	4	21.1	1021 10	049548	RETROTRANSPOSON - LIKE	2.02e+03	947	4	21.1	2109 14	086125	237K PROTEIN.	2.02e+03
875	4	21.1	1023 14	092787	DNA POLYMERASE.	2.02e+03	948	4	21.1	2161 14	091901	REPLICASE.	2.02e+03
876	4	21.1	1028 11	007409	PLASMACYTOMA-ASSOCIATE	2.02e+03	949	4	21.1	2161 14	091724	RNA-DEPENDENT RNA POLY	2.02e+03
877	4	21.1	1028 11	062682	BIG-1 PROTEIN PRECURSO	2.02e+03	950	4	21.1	2162 14	091940	RNA-DEPENDENT RNA POLY	2.02e+03
878	4	21.1	1048 3	P78981	NA+-ATPASE.	2.02e+03	951	4	21.1	2165 14	041355	POLYMERASE SUBUNIT L.	2.02e+03
879	4	21.1	1059 14	076630	POL POLYPROTEIN (FRAGM	2.02e+03	952	4	21.1	2165 14	041356	POLYMERASE SUBUNIT L.	2.02e+03
880	4	21.1	1061 14	041065	PECV-1 DNA TOPOISOMERA	2.02e+03	953	4	21.1	2165 14	009721	POLYMERASE (L).	2.02e+03
881	4	21.1	1064 5	025773	ASPARAGINE-RICH ANTIGE	2.02e+03	954	4	21.1	2183 14	064962	VIRAL RNA-DEPENDENT RN	2.02e+03
882	4	21.1	1074 14	065817	U52.	2.02e+03	955	4	21.1	2191 14	066474	POLYPROTEIN.	2.02e+03
883	4	21.1	1095 11	060784	FIBRILLIN-1 (FRAGMENT)	2.02e+03	956	4	21.1	2204 14	090341	LARGE POLYMERASE PROTE	2.02e+03
884	4	21.1	1103 14	041928	SSDNA BINDING PROTEIN.	2.02e+03	957	4	21.1	2218 14	067817	COMPLETE GENOME.	2.02e+03
885	4	21.1	1116 2	067838	HYPOTHETICAL 127.8 KD	2.02e+03	958	4	21.1	2228 14	055528	RNA POLYMERASE PROTEIN	2.02e+03
886	4	21.1	1123 14	066933	POLYMERASE (FRAGMENT).	2.02e+03	959	4	21.1	2233 14	081890	L PROTEIN.	2.02e+03
887	4	21.1	1128 1	051999	REPI PROTEIN.	2.02e+03	960	4	21.1	2254 11	054898	LOW VOLTAGE-ACTIVATED.	2.02e+03
888	4	21.1	1134 1	030078	MOLYBDOPHTERIN OXIDORE	2.02e+03	961	4	21.1	2258 14	090245	POLYPROTEIN PRECURSOR.	2.02e+03
889	4	21.1	1136 3	043044	HYPOTHETICAL 129.8 KD	2.02e+03	962	4	21.1	2258 14	093128	RNAI POLYPROTEIN.	2.02e+03
890	4	21.1	1138 14	065716	ALPHA PROTEIN.	2.02e+03	963	4	21.1	2258 14	081081	RNAI POLYMERASE.	2.02e+03
891	4	21.1	1144 11	070515	NITRIC OXIDE SYNTHASE	2.02e+03	964	4	21.1	2258 14	089238	LARGE PROTEIN.	2.02e+03
892	4	21.1	1146 14	089468	POLYMERASE POLYPROTEIN	2.02e+03	965	4	21.1	2292 14	066850	COMPLETE RNA GENOME PR	2.02e+03
893	4	21.1	1146 14	089472	POLYMERASE POLYPROTEIN	2.02e+03	966	4	21.1	2334 14	096725	RNA.	2.02e+03
894	4	21.1	1156 14	093209	POL PROTEIN.	2.02e+03	967	4	21.1	2352 5	061240	HRNOTCH PROTEIN.	2.02e+03
895	4	21.1	1159 2	011028	PROBABLE TRANSCRIPTION	2.02e+03	968	4	21.1	2391 5	027732	CARBAMOYL-PHOSPHATE SY	2.02e+03
896	4	21.1	1170 13	057612	JAK1 TYROSINE KINASE.	2.02e+03	969	4	21.1	2470 13	090681	CATION- INDEPENDENT MAN	2.02e+03



```

970 4 21.1 2842 14 093197 POLYPROTEIN. 2.02e+03
971 4 21.1 2842 14 089252 POLYPROTEIN. 2.02e+03
972 4 21.1 2843 14 089251 POLYPROTEIN. 2.02e+03
973 4 21.1 2873 14 P89967 RNA FOR POLYPROTEIN, C 2.02e+03
974 4 21.1 2873 14 090481 POLYPROTEIN. 2.02e+03
975 4 21.1 2894 1 058791 HYPOTHETICAL PROTEIN M 2.02e+03
976 4 21.1 2967 14 041892 POLYPROTEIN PRECURSOR. 2.02e+03
977 4 21.1 3005 14 056074 POLYPROTEIN. 2.02e+03
978 4 21.1 3035 14 072347 POLYPROTEIN. 2.02e+03
979 4 21.1 3059 14 085197 POLYPROTEIN. 2.02e+03
980 4 21.1 3061 14 085265 POLYPROTEIN. 2.02e+03
981 4 21.1 3072 14 092645 POLYPROTEIN. 2.02e+03
982 4 21.1 3086 14 089525 POLYPROTEIN. 2.02e+03
983 4 21.1 3141 14 P89038 RNA GENOME. 2.02e+03
984 4 21.1 3199 14 085074 POLYPROTEIN. 2.02e+03
985 4 21.1 3222 14 085071 POLYPROTEIN. 2.02e+03
986 4 21.1 3222 14 085057 POLYPROTEIN. 2.02e+03
987 4 21.1 3391 14 P89313 POLYPROTEIN. 2.02e+03
988 4 21.1 3391 14 P89315 POLYPROTEIN. 2.02e+03
989 4 21.1 3391 14 P89314 POLYPROTEIN. 2.02e+03
990 4 21.1 3473 14 083034 POLYPROTEIN. 2.02e+03
991 4 21.1 3493 14 039734 POLYPROTEIN. 2.02e+03
992 4 21.1 3534 14 039266 COUNTERPART OF HSV-1 G 2.02e+03
993 4 21.1 3588 2 060070 LICHENYSLIN SYNTHETASE 2.02e+03
994 4 21.1 3722 2 P94873 ALPHA-AMINOADIPYL-CYST 2.02e+03
995 4 21.1 3729 2 033956 TYLACTONE SYNTHASE MOD 2.02e+03
996 4 21.1 3898 14 092365 POLYPROTEIN. 2.02e+03
997 4 21.1 3898 14 092872 POLYPROTEIN. 2.02e+03
998 4 21.1 4017 14 088508 (PURDUE-115). 2.02e+03
999 4 21.1 4957 4 014687 ALR. 2.02e+03
1000 4 21.1 26926 4 010466 TITIN, HEART ISOFORM N 2.02e+03

```

## ALIGNMENTS

```

RESULT 1
ID O88826 PRELIMINARY: PRT: 913 AA.
AC O88826;
DT 01-NOV-1998 (TREMBREL. 08, CREATED)
DT 01-NOV-1998 (TREMBREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE GOB-5 PROTEIN.
GN GOB-5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCITUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-INTESTINE;
RA KOMIYA T., TANIGAWA Y., HIROHASHI S.;
RT "Cloning of a gene, gob-5, which is expressed in intestinal goblet
RT cells in mice."
RL SUBMITTED (AUG-1998) TO EMBL/GENBANK/DDBJ DATA BANKS.
DR EMBL; AB017136; D1034712; -.
SQ SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;

Query Match 52.6%; Score 10; DB 11; Length 913;
Best Local Similarity 100.0%; Pred. No. 3.73e-09;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 169 YNNDKFKYLS 178
| | | | | | | | | |
QY 1 YNNDKFKYLS 10

RESULT 2
ID O45500 PRELIMINARY: PRT: 433 AA.
AC O45500;
DT 01-JUN-1998 (TREMBREL. 06, CREATED)
DT 01-JUN-1998 (TREMBREL. 06, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBREL. 09, LAST ANNOTATION UPDATE)
DE F39E2.8 PROTEIN.
GN F39E2.8.

```

```

OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RA DOBSON R.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL NATURE 368:32-38(1994).
DR EMBL; Z92834; E1348581; -.
SQ SEQUENCE 433 AA; 49673 MW; D47D3C20 CRC32;

Query Match 36.8%; Score 7; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.17e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 284 KFYLSNG 290
| | | | | | | | | |
QY 6 KFYLSNG 12

```

```

RESULT 3
ID O27930 PRELIMINARY: PRT: 97 AA.
AC O27930;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBREL. 07, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 10.4 KD PROTEIN.
GN MTH1908.
OS METHANOBACTERIUM THERMOAUTOTROPHICUM.
OC ARCHAEA; EURYARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;
OC METHANOBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics."
RL J. BACTERIOL. 179:7135-7155(1997).
DR EMBL; AE000942; G2623042; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 97 AA; 10432 MW; 7741311C CRC32;

Query Match 31.6%; Score 6; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.02e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 23 FYLSNG 28
| | | | | | | | | |
QY 7 FYLSNG 12

```



```

RESULT 4
ID O16453 PRELIMINARY; PRT: 170 AA.
AC O16453:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE C54F6.14 PROTEIN.
GN C54F6.14.
OS CAENORHABDITIS ELEGANS.
OC EURARYOTA; METAZOA; NEMATODA; SECERNENTEA; RHABDITIA; RHABDITIDA;
OC RHABDITINA; RHABDITOIDEA; RHABDITIDAE; PELODERINAE; CAENORHABDITIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE; 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P., III of C.
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL NATURE 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX GEISEL C., BRADSHAW H.;
RA SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA WATERSTON R.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR PFAM; AF016447; G2315445; -.
DR PFAM; PF00210; ferritin; 1.
SQ SEQUENCE 170 AA; 19525 MW; 6479E209 CRC32;

Query Match 31.6%; Score 6; DB 5; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.02e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 67 RIOAVER 72
QY 13 RIOAVER 18

RESULT 5
ID P75163 PRELIMINARY; PRT: 181 AA.
AC P75163:
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE C12.ORI1810 PROTEIN.
OS MYCOPLASMA PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M129;
RX MEDLINE; 97105885.
RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.C.,
RA HERRMANN R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M129;

```

```

RA HIMMELREICH R., HILBERT H., LI B.C.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000021; G1673873; -.
SQ SEQUENCE 181 AA; 21358 MW; CC6FFED1 CRC32;

Query Match 31.6%; Score 6; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 1.02e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 125 SNGRIQ 130
QY 10 SNGRIQ 15

RESULT 6
ID Q55493 PRELIMINARY; PRT: 314 AA.
AC Q55493:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE AUXIN-INDUCED PROTEIN.
OS SYNECHOCYSTIS SP. (STRAIN PCC 6803).
OC BACTERIA; CYANOBACTERIA; CHROCOCCALES; SYNECHOCYSTIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RA TABATA S.;
RL SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE; 96127529.
RA KANEKO T., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA SUGIURA M., TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. sequence features in the 1Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA RES. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE; 97061201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAKOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NAROO K., OKUMURA S.,
RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. PCC6803. II. Sequence determination of the entire
RT genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
DR EMBL; D64006; D1011493; -.
DR PFAM; PF00248; aldo_ket_red; 2.
SQ SEQUENCE 314 AA; 34632 MW; 11D9F821 CRC32;

Query Match 31.6%; Score 6; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.02e+00;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 144 GRIQAV 149
QY 12 GRIQAV 17

RESULT 7
ID Q48999 PRELIMINARY; PRT: 316 AA.
AC Q48999:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE SIMILAR TO ISOMALTASE-SUCRASE (FRAGMENT).
OS MYCOPLASMA CAPRICOLUM.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLICUTES;

```

OC CAPRICOLUM GROUP.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 27343(KID);  
RX MEDLINE; 96059641.  
RA BORK P., OZOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,  
RA GILBERT W., GILLEVEY P.M.;  
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
RT its physiology."  
RL MOL. MICROBIOL. 16:955-967(1995).  
DR EMBL; Z33074; G530423; -.  
FT .NON\_TER 316 316  
SQ SEQUENCE 316 AA; 37025 MW; 2AF7493F CRC32;

Query Match 31.6%; Score 6; DB 2; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 173 NDERFY 178  
QY 3 NDERFY 8

RESULT 8  
ID Q38909 PRELIMINARY; PRT; 332 AA.  
AC Q38909; 004750;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE XYLOGLUCAN ENDOTRANSGLYCOSYLASE-RELATED PROTEIN.  
GN XTR2 OR EXGT-A2.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA: TRACHEOPHYTA;  
OC EUPHYLLOPHYTES: SPERMATOPHYTA: MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE; 96267577.  
RA XU W., CAMPBELL P., VARGHESE A.K., BRAAM J.;  
RT "The Arabidopsis XET-related gene family: environmental and hormonal  
RT regulation of expression."  
RL PLANT J. 9:879-889(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA OKAWOTO S., KAMIMAI T., NISHITANI K.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U43487; G1244756; -.  
DR EMBL; D63510; D1021121; -.  
DR PFAM; PF00722; glycosyl\_hydro9; 1.  
DR MENDEL; 6227; Arabid; 1017.6.  
KW TRANSFERASE.  
SQ SEQUENCE 332 AA; 38252 MW; A3BC9553 CRC32;

Query Match 31.6%; Score 6; DB 10; Length 332;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 FYLSNG 99  
QY 7 FYLSNG 12

RESULT 9  
ID Q49982 PRELIMINARY; PRT; 357 AA.  
AC Q49982;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE FLAVANONE 3-HYDROXYLASE.  
OS DENDRATHEMA X GRANDIFLORUM.  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;

OC CAPRICOLUM GROUP.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 27343(KID);  
RX MEDLINE; 96059641.  
RA BORK P., OZOUNIS C., CASARI G., SCHNEIDER R., SANDER C., DOLAN M.,  
RA GILBERT W., GILLEVEY P.M.;  
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals  
RT its physiology."  
RL MOL. MICROBIOL. 16:955-967(1995).  
DR EMBL; Z33074; G530423; -.  
FT .NON\_TER 316 316  
SQ SEQUENCE 316 AA; 37025 MW; 2AF7493F CRC32;

Query Match 31.6%; Score 6; DB 2; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 173 NDERFY 178  
QY 3 NDERFY 8

RESULT 8  
ID Q38909 PRELIMINARY; PRT; 332 AA.  
AC Q38909; 004750;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE XYLOGLUCAN ENDOTRANSGLYCOSYLASE-RELATED PROTEIN.  
GN XTR2 OR EXGT-A2.  
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA: TRACHEOPHYTA;  
OC EUPHYLLOPHYTES: SPERMATOPHYTA: MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;  
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE; 96267577.  
RA XU W., CAMPBELL P., VARGHESE A.K., BRAAM J.;  
RT "The Arabidopsis XET-related gene family: environmental and hormonal  
RT regulation of expression."  
RL PLANT J. 9:879-889(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA OKAWOTO S., KAMIMAI T., NISHITANI K.;  
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; U43487; G1244756; -.  
DR EMBL; D63510; D1021121; -.  
DR PFAM; PF00722; glycosyl\_hydro9; 1.  
DR MENDEL; 6227; Arabid; 1017.6.  
KW TRANSFERASE.  
SQ SEQUENCE 332 AA; 38252 MW; A3BC9553 CRC32;

Query Match 31.6%; Score 6; DB 10; Length 332;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 94 FYLSNG 99  
QY 7 FYLSNG 12

RESULT 9  
ID Q49982 PRELIMINARY; PRT; 357 AA.  
AC Q49982;  
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
DE FLAVANONE 3-HYDROXYLASE.  
OS DENDRATHEMA X GRANDIFLORUM.  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;

OC EUPHYLLOPHYTES: SPERMATOPHYTA: MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC ASTERIDAE; ASTERALES; ASTERACEAE; DENDRATHEMA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PARK S.Y., KIM Y.H.;  
RL PLANT PHYSIOL. 115:1731-1731(1997).  
DR EMBL; U86837; G2801407; -.  
SQ SEQUENCE 357 AA; 40353 MW; B17725D7 CRC32;

Query Match 31.6%; Score 6; DB 10; Length 357;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 263 YLSNGR 268  
QY 8 YLSNGR 13

RESULT 10  
ID Q04112 PRELIMINARY; PRT; 372 AA.  
AC Q04112;  
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)  
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE FLAVANONE 3-HYDROXYLASE.  
OS PERILLA FRUTESCENS.  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES: SPERMATOPHYTA: MAGNOLIOPHYTA; EUDICOTYLEDONS;  
OC ASTERIDAE; GENTIANANAE; LAMIALES; LAMIACEAE; PERILLA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 98088016.  
RA GONG Z., YAMAZAKI M., SUGIYAMA M., TANAKA Y., SAITO K.;  
RT "Cloning and molecular analysis of structural genes involved in  
RT anthocyanin biosynthesis and expressed in a forma-specific manner in  
RT Perilla frutescens."  
RL PLANT MOL. BIOL. 35:915-927(1997).  
DR EMBL; AB002816; D1020434; -.  
DR PFAM; PF00671; FeAsc.Oxidore; 1.  
DR MENDEL; 15157; Perfr1186; mal15157;  
SQ SEQUENCE 372 AA; 41948 MW; E44844A7 CRC32;

Query Match 31.6%; Score 6; DB 10; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 267 YLSNGR 272  
QY 8 YLSNGR 13

RESULT 11  
ID Q43741 PRELIMINARY; PRT; 374 AA.  
AC Q43741;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
DE NARINGENIN 3-DIOXYGENASE (EC 1.14.11.9)  
DE (NARINGENIN/2-OXOGLUTARATE 3-DIOXYGENASE) (FLAVANONE 3-HYDROXYLASE).  
GN FHT.  
OS BROMHEADIA FINLAYSONIANA.  
OC EUKARYOTA: VIRIDIPLANTAE: STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
OC EUPHYLLOPHYTES: SPERMATOPHYTA; MAGNOLIOPHYTA; LILLOPSIDA; ASPARAGALES;  
OC ORCHIDACEAE; BROMHEADIA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. ORCHID; TISSUE=FLOWER;  
RA LIM S.H.;  
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- CATALYTIC ACTIVITY: NARINGENIN + 2-OXOGLUTARATE + O(2) -  
CC DIHYDROKAEMPFEROL + SUCCINATE + CO(2).  
CC -1- COFACTOR: IRON; ASCORBATE.  
DR EMBL; X89199; G895743; -.

DR PFAM: PF00671; Fe\_Asc\_oxidored; 1.  
 DR MENDEL: 11552; BROfi.1186;1.  
 KW DIOXYGENASE; OXIDOREDUCTASE.  
 SQ SEQUENCE 374 AA; 41733 MW; A72CA068 CRC32;

Query Match 31.6%; Score 6; DB 10; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 266 YLSNGR 271  
 |||||  
 QY 8 YLSNGR 13

RESULT 12  
 ID O15723 PRELIMINARY; PRT: 382 AA.  
 AC O15723;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE LTR-RETROTRANSPON SKIPPER, GAG (GAG).  
 GN GAG.

OS DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).  
 OC EUKARYOTA; DICTYOSTELIIDA; DICTYOSTELIUM.

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=AX3, AX2;  
 RX MEDLINE: 98186844.

RA LENG P., KLAETTE D.H., SCHUMANN G., BOEKE J.D., STECK T.L.;  
 RT "Skipper, an LTR retrotransposon of Dictyostelium.";

RL NUCLEIC ACIDS RES. 26:2008-2015(1998).

DR EMBL: AF017040; G2367673; -.

DR EMBL: AF049230; G3135477; -.

SQ SEQUENCE 382 AA; 42669 MW; 9068B5E8 CRC32;

Query Match 31.6%; Score 6; DB 5; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 286 NNDKX 291  
 |||||  
 QY 2 NNDKX 7

RESULT 13  
 ID O24679 PRELIMINARY; PRT: 410 AA.  
 AC O24679;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE TECA4.

GN TECA4.

OS BURKHOLDERIA SP.

OC BACTERIA; PROTEOBACTERIA; BETA SUBDIVISION; BURKHOLDERIA GROUP;  
 OC BURKHOLDERIA.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PS12;

RX MEDLINE: 97390087.

RA BEIL S., HAPPE B., TIMMIS K.N., PIEPER D.H.;

RT "Genetic and biochemical characterization of the broad spectrum  
 chlorobenzene dioxygenase from Burkholderia sp. strain

PS12--dechlorination of 1,2,4,5-tetrachlorobenzene.";

RL EUR. J. BIOCHEM. 247:190-199(1997).

DR EMBL: U78099; G2264420; -.

SQ SEQUENCE 410 AA; 42961 MW; 0A9A208A CRC32;

Query Match 31.6%; Score 6; DB 2; Length 410;  
 Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 359 GRIQAV 364  
 |||||

QY 12 GRIQAV 17  
 RESULT 14  
 ID Q52386 PRELIMINARY; PRT: 410 AA.  
 AC Q52386;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE CHLOROBENZENE DIOXYGENASE, NADH-FERREDOXIN REDUCTASE.  
 GN TCBA4.

OS PSEUDOMONAS SP.

OG PLASMID PP51.

OC BACTERIA; PROTEOBACTERIA.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 96223965.

RA WERLEN C., KOHLER H.E., DER MEER J.R.;

RT "The broad substrate chlorobenzene dioxygenase and cis-chlorobenzene  
 dihydrodiol dehydrogenase of Pseudomonas sp. strain P51 are linked  
 evolutionarily to the enzymes for benzene and toluene degradation.";

RL J. BIOL. CHEM. 271:4009-4016(1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 91100319.

RA DER MEER J.R., VAN NEERVEN A.R., DE VRIES E.J., DE VOS W.M.,

RA ZEHNDER A.J.;

RT "Cloning and characterization of plasmid-encoded genes for the  
 degradation of 1,2-dichloro-, 1,4-dichloro-, and

RT 1,2,4-trichlorobenzene of Pseudomonas sp. strain P51.";

RL J. BACTERIOL. 173:6-15(1991).

DR EMBL: U15298; G557073; -.

KW DIOXYGENASE; PLASMID.

SQ SEQUENCE 410 AA; 43084 MW; 3BD15470 CRC32;

Query Match 31.6%; Score 6; DB 2; Length 410;  
 Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 359 GRIQAV 364  
 |||||  
 QY 12 GRIQAV 17

RESULT 15  
 ID O26298 PRELIMINARY; PRT: 458 AA.  
 AC O26298;  
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)  
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)  
 DE CONSERVED PROTEIN.  
 GN MTH196.

OS METHANOBACTERIUM THERMOAUTOTROPHICUM.

OC ARCHAEA; EURARCHAEOTA; METHANOBACTERIALES; METHANOBACTERIACEAE;

OC METHANOBACTERIUM.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DELTA H;

RX MEDLINE: 98037514.

RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,

RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,

RA HARRISON D., HOANG L., KEAGLE P., LOMW W., POTIER B., QIU D.,

RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,

RA JIWANI N., CARUSO A., BUSH D., SAFFER H., PATWELL D., PRASHAKAR S.,

RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,

RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;

RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 deltaH: functional analysis and comparative genomics.";

RL J. BACTERIOL. 179:7135-7155(1997).

DR EMBL: AE000807; G2621241; -.

SQ SEQUENCE 458 AA; 53966 MW; 64DADC9D CRC32;

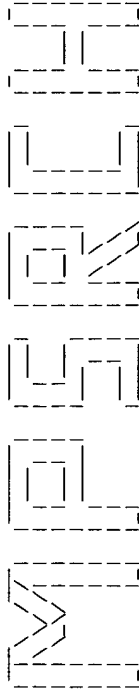
Query Match 31.6%; Score 6; DB 1; Length 458;

Best Local Similarity 100.0%; Pred. No. 1.02e+00;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 129 LSNPRI 134  
      |||||  
Qy 9 LSNPRI 14

Search completed: Sat Aug 28 14:38:54 1999  
Job time : 48 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 23 13:06:45 1999; MasPar time 15.63 Seconds  
32.658 Million cell updates/sec

Tabular output not generated.

Title: >US-09-049-696-43  
Description: (1-24) from US09049696.pcp  
Perfect Score: 189  
Sequence: 1 YTKRCTFNKVTGLYEKGCFVLQS 24

Scoring table: PAM 150  
Gap 15

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq35  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29 30:part30 31:part31 32:part32 33:part33  
34:part34 35:part35 36:part36 37:part37 38:part38  
39:part39

Statistics: Mean 21.166; Variance 65.930; scale 0.321

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	64	33.9	246	35	Human amyloid beta-pr	2.28e+01
2	63	33.3	165	22	H. pylori cytoplasmic	2.89e+01
3	63	33.3	420	13	BAV3 E1B 56K protein.	2.89e+01
4	62	32.8	396	19	Interleukin-2 recepto	3.55e+01
5	61	32.3	345	22	H. pylori cytoplasmic	4.61e+01
6	60	31.7	93	33	Nucleotide sequence o	5.80e+01
7	60	31.7	180	33	Amino acid sequence o	5.80e+01
8	60	31.7	491	30	Hiv-1 breakthrough is	5.80e+01
9	60	31.7	1404	7	Sequence of a serrate	5.80e+01
10	59	31.2	642	28	Yeast acyl-coenzyme A	7.30e+01
11	59	31.2	642	29	Yeast acyl-coenzyme A	7.30e+01
12	58	30.7	28	18	ME#1 library derived	9.18e+01
13	58	30.7	255	2	Sequence of a region	9.18e+01
14	58	30.7	306	35	Helicobacter polyepit	9.18e+01
15	58	30.7	388	19	Rat superior cervical	9.18e+01
16	58	30.7	536	2	Sequence of a region	9.18e+01

17	58	30.7	735	35	W69761	Acetobacter xylinum b	9.18e+01
18	58	30.7	855	10	R58385	FIV UK8 envelope prot	9.18e+01
19	58	30.7	870	10	R51246	Composite sequence of	9.18e+01
20	57	30.2	109	28	W47068	Human brain P2X-1 rec	1.15e+02
21	57	30.2	114	38	W90000	Expressed antigen for	1.15e+02
22	57	30.2	250	38	W89882	Antigen 1 from cluste	1.15e+02
23	57	30.2	250	38	W89998	Expressed antigen for	1.15e+02
24	57	30.2	250	29	W55488	H. pylori ORF hp5e152	1.15e+02
25	57	30.2	250	34	W62864	Helicobacter pylori o	1.15e+02
26	57	30.2	250	35	W71528	Helicobacter polyepit	1.15e+02
27	57	30.2	269	29	W55341	H. pylori ORF hp5e152	1.15e+02
28	57	30.2	338	38	W89928	Antigen 1 from cluste	1.15e+02
29	57	30.2	353	38	W89824	Protein encoded by cl	1.15e+02
30	57	30.2	388	28	W47066	Human brain P2X-1 rec	1.15e+02
31	57	30.2	388	33	W55035	HPURR amino acid sequ	1.15e+02
32	57	30.2	396	19	R97570	Interleukin-2 recepto	1.15e+02
33	57	30.2	404	28	W47059	Human brain P2X-2 rec	1.15e+02
34	57	30.2	496	30	W37059	HIV-1 breakthrough is	1.15e+02
35	57	30.2	496	30	W37058	HIV-1 breakthrough is	1.15e+02
36	57	30.2	551	39	W73480	Grapevine leafroll vi	1.15e+02
37	56	29.6	224	10	R53637	Bovine trypsin.	1.44e+02
38	56	29.6	335	16	R88258	Eucalyptus gunnii cin	1.44e+02
39	56	29.6	335	26	W36046	Eucalyptus cinnamomyl	1.44e+02
40	56	29.6	616	13	R73007	Aminopeptidase 012 cl	1.44e+02
41	56	29.6	1385	14	R44201	Bacillus thuringiensis	1.44e+02
42	56	29.6	1385	3	R20066	B.thuringiensis toxin	1.44e+02
43	56	29.6	1385	21	W13884	17a toxin.	1.44e+02
44	56	29.6	1385	6	R28889	Toxin 17a.	1.44e+02
45	56	29.6	1385	6	R29516	BT toxin 17a.	1.44e+02

ALIGNMENTS

RESULT 1  
ID W64260 standard; Protein; 246 AA.  
AC W64260;  
DT 24-NOV-1998 (first entry)  
DE Human amyloid beta-protein precursor inhibitor.  
KW MCP-7; mast cell protease 7; trypsinase-7; serine protease;  
KW amyloid beta-protein precursor inhibitor; human; blood clot;  
KW anticoagulant; myocardial infarction; reocclusion; thromboembolism;  
KW cerebral embolism; thrombosis; therapy.  
OS Homo sapiens.  
PN W09824886-A1.  
PD 11-JUN-1998.  
PF 25-NOV-1997; U21620.  
PR 04-DEC-1996; US-092354.  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
PI Stevens RL;  
DR WPI; 98-333308/29.  
PT New compositions containing trypsinase-7, e.g. mouse mast cell  
PT protease-7 - are used to treat clot formation in e.g. myocardial  
PT infarction, reocclusion following angioplasty or pulmonary  
PT thromboembolism  
PS Example: Page 77; 92pp; English.  
CC This polypeptide comprises the human amyloid beta-protein precursor  
CC inhibitor (ABPI). The crystallographic structure of the complex of  
CC rat trypsin with ABPI was used as a template structure to model the  
CC structure of the substrate-binding pocket of mouse mast cell  
CC protease 7 (mMCP-7, see W64233). The invention relates to mMCP-7  
CC and related trypsinase-7 proteases that can be used to prevent or  
CC inhibit fibrin clot formation. Such proteases can be used to treat  
CC disorders mediated by undesirable thrombus clot formation such as  
CC myocardial infarction and reocclusion following angioplasty of  
CC blood clots associated with pulmonary thromboembolism, deep vein  
CC thrombosis, cerebral embolism, renal vein and peripheral arterial  
CC thrombosis.  
SQ Sequence 246 AA;  
Query Match 33.9%; Score 64; DB 35; Length 246;  
Best Local Similarity 41.2%; Pred. No. 2.28e+01;  
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;



PS Claim 61: Page 1169-1170; 1481pp; English.  
 CC The present sequence is a Helicobacter pylori cytoplasmic protein.  
 CC The protein may be used in a vaccine to prevent or treat H. pylori  
 CC infection or to identify H. pylori polypeptide binding compounds, useful  
 CC as potential H. pylori life cycle activators or inhibitors. The genomic  
 CC sequence of H. pylori (ATCC 55679) was determined from overlapping  
 CC contigs generated by mechanically shearing the bacterial DNA. The  
 CC sequences were analysed for ORF of at least 180 nucleotides, and the  
 CC predicted coding regions defined by computer evaluation. To identify  
 CC likely H. pylori antigens for vaccine development, the amino acid  
 CC sequences predicted from various ORF were analysed for significant  
 CC homology to other known or exported membrane proteins. Having identified  
 CC and determined the sequences of interest, particular regions can be  
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide  
 CC production, e.g. in E. coli hosts.  
 CC Sequence 345 AA;  
 SQ

Query Match 32.3%; Score 61; DB 22; Length 345;  
 Best Local Similarity 47.1%; Pred. No. 4.61e+01;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 188 ftksnfpvialyerg 204  
 :|| | | :|||:  
 QY 1 YTRCTFNKVTGLYKRG 17

RESULT 6 standard; Protein; 93 AA.  
 ID W59779 standard; Protein; 93 AA.

AC W59779;  
 DT 12-OCT-1998 (first entry)  
 DE Nucleotide sequence of complete EST3 gene product 1.  
 KW EST3 gene: telomerase inhibitor; cancer; mammal; tumour growth;  
 KW stem cell; chemotherapeutic agent.  
 OS Saccharomyces cerevisiae.

PN W09823759-A2.  
 PD 04-JUN-1998.  
 PF 26-NOV-1997; U21272.  
 PR 26-NOV-1996; US-756693.  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 PI Lundblad V;  
 DR WPI: 98-332927/29.  
 DR N-PSDB: V41581.

PT Telomerase-associated proteins useful to identify telomerase  
 PT inhibitors - useful e.g. for cancer treatment in mammals, and  
 PT screening methods to isolate additional telomerase-associated genes  
 PT or protein homologues

PS Disclosure; Fig 7; 71pp; English.  
 CC This is the nucleotide sequence of the EST3 gene product 1 (including  
 CC flanking regions and the +1 ribosomal frame shift), used in the method  
 CC of the invention to identify telomerase inhibitors which are useful in  
 CC the treatment of cancer. The proteins are useful to isolate  
 CC telomerase-inhibiting compounds. Such inhibitors are useful in cancer  
 CC treatment in mammals, since reactivation of telomerase (normally present  
 CC only in germ line cells) is thought to be necessary for sustained tumour  
 CC growth, and only tumour and stem cells would be targeted by such agents,  
 CC producing limited side effects compared to chemotherapeutic agents.  
 CC Sequence 93 AA;  
 SQ

Query Match 31.7%; Score 60; DB 33; Length 93;  
 Best Local Similarity 36.8%; Pred. No. 5.80e+01;  
 Matches 7; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Db 61 tnpchfakitkfyv-vcdy 78  
 :| | | | :| | | |:  
 QY 2 TKRCTFNKVTGLYKRGCE 20

RESULT 7 standard; protein; 180 AA.  
 ID W59797 standard; protein; 180 AA.

AC W59797;  
 DT 12-OCT-1998 (first entry)  
 DE Amino acid sequence of EST3 gene product.  
 KW EST3 gene product; telomerase inhibitor; cancer; mammal; stem cell;

KW tumour growth; chemotherapeutic agent.  
 OS Saccharomyces cerevisiae.

PN W09823759-A2.  
 PD 04-JUN-1998.  
 PF 26-NOV-1997; U21272.  
 PR 26-NOV-1996; US-756693.  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 PI Lundblad V;  
 DR WPI: 98-332927/29.

PT Telomerase-associated proteins useful to identify telomerase  
 PT inhibitors - useful e.g. for cancer treatment in mammals, and  
 PT screening methods to isolate additional telomerase-associated genes  
 PT or protein homologues

PS Claim 23; Fig 6; 71pp; English.  
 CC This is the amino acid sequence of the EST3 gene product used in the  
 CC method of the invention to identify telomerase inhibitors which are  
 CC useful in the treatment of cancer. The proteins are useful to isolate  
 CC telomerase-inhibiting compounds. Such inhibitors are useful in cancer  
 CC treatment in mammals, since reactivation of telomerase (normally present  
 CC only in germ line cells) is thought to be necessary for sustained tumour  
 CC growth, and only tumour and stem cells would be targeted by such agents,  
 CC producing limited side effects compared to chemotherapeutic agents.  
 CC Sequence 180 AA;  
 SQ

Query Match 31.7%; Score 60; DB 33; Length 180;  
 Best Local Similarity 36.8%; Pred. No. 5.80e+01;  
 Matches 7; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

Db 61 tnpchfakitkfyv-vcdy 78  
 :| | | | :| | | |:  
 QY 2 TKRCTFNKVTGLYKRGCE 20

RESULT 8

ID W37057 standard; Protein; 491 AA.

AC W37057;  
 DT 20-JUL-1998 (first entry)  
 DE HIV-1 breakthrough isolate clone C8.6 gp120 polypeptide.  
 KW HIV-1; envelope protein; gp120; MN-rgp120; vaccine; AIDS.  
 OS Human immunodeficiency virus type 1.  
 PN W09801564-A1.  
 PD 15-JAN-1998.

PF 03-JUL-1997; U09690.  
 PR 08-JUL-1996; US-676737.  
 PA (GETH ) GENENTECH INC.  
 PI Berman PW;

PT Human immunodeficiency virus gp120 sequences from vaccine  
 PT breakthrough strains - useful in providing added protection against  
 PT HIV not provided by original vaccine  
 PS Claim 1; Page 104-106; 193pp; English.

CC This is the deduced sequence of a gp120 envelope polypeptide of  
 CC clone C8.6 of a HIV-1 breakthrough isolate obtained from an  
 CC individual infected with HIV-1 through high risk activity while  
 CC participating in Phase I or Phase 2 trials of MN-rgp120, a  
 CC candidate recombinant gp120 HIV-1 vaccine. Nucleotide sequences  
 CC (see V00511-30) for gp120 polypeptides (see W37054-67) were  
 CC obtained from 2 clones of each of 7 breakthrough isolates. All 7  
 CC envelope glycoproteins have sequences typical of subtype (clade) B  
 CC viruses. The overall homology with MN-rgp120 is 69-80%. Use of  
 CC the gp120 polypeptides from one or more of the breakthrough  
 CC isolates, usually together with MN-rgp120, can provide protection  
 CC against HIV strains that are sufficiently different from the  
 CC vaccine strain (e.g. MN-rgp120) that the vaccine does not confer  
 CC protection against those strains. The gp120 proteins can also be  
 CC used in screening assays to identify antagonists of CC-CR chemokine  
 CC receptors. Antibodies induced by the polypeptides are also provided.  
 CC Sequence 491 AA;  
 SQ

Query Match 31.7%; Score 60; DB 30; Length 491;  
 Best Local Similarity 53.3%; Pred. No. 5.80e+01;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 118 kncsfnittslrdkg 132  
 ||||| :|| :||  
 QY 3 KRCFTNKVTGLYEKG 17

## RESULT 9

ID R38304 standard; Protein: 1404 AA.  
 AC R38304;  
 DT 30-NOV-1993 (first entry)  
 DE Sequence of a serrate protein.  
 KW Serrate; topoxythmic protein; family.  
 OS Drosophila melanogaster.  
 PN WO9312141-A.  
 PD 24-JUN-1993.  
 PF 11-DEC-1991; U09240.  
 PR 11-DEC-1991; WO-U09240.  
 PA (UYVA ) UNIV VALE.  
 PI Artavanis-Tsakonas S, Fleming RJ;  
 DR WPI: 93-214095/26.  
 DR N-PSDB: 043910.  
 PT Purified serrate protein, nucleic acid and antibodies - used in  
 PT the study and manipulation of differentiation and other  
 PT physiological processes  
 PS Claim 4; Pages 74-80; 119pp; English.  
 CC Two Drosophila genomic phage libraries were screened and recombinant  
 CC clones were isolated. The cDNAs in lambda gt10 were isolated from an  
 CC early pupal library. The C1 cDNA was isolated from an early pupal  
 CC library. Subsequently the C3 cDNA was isolated using the 5', 700 bp  
 CC terminal fragment of the C1 cDNA as probe. The complete 5561bp  
 CC sequence of DNA of the Drosophila Serrate protein was derived from  
 CC C1 and C3 cDNAs (043910). The deduced protein product appears to be  
 CC a transmembrane protein. AAs 51-80 represent the likely signal  
 CC peptide; aas 542-564 represent potential membrane associated region;  
 CC aas 1221-1245 represent the putative transmembrane domain.  
 SQ Sequence 1404 AA;

Query Match 31.7%; Score 60; DB 7; Length 1404;  
 Best Local Similarity 38.1%; Pred. No. 5.80e+01;  
 Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Db 141 srccsfgnattkllggssfv1 161  
 :|||:| :|||  
 QY 2 TKRCFTNKVTGLYEKGCEFLV 22

## RESULT 10

ID W38418 standard; Protein: 642 AA.  
 AC W38418;  
 DT 21-MAY-1998 (first entry)  
 DE Yeast acyl-coenzyme A:cholesterol acyltransferase 2.  
 KW Acyl-coenzyme A: cholesterol acyltransferase 2; ARE2; yeast;  
 KW sterol; esterification; arteriosclerosis; hyperlipidaemia;  
 KW antifungal; fungicide.  
 OS Saccharomyces cerevisiae.  
 FH Key Location/Qualifiers  
 FT Domain 215..231  
 FT /note= "putative transmembrane domain"  
 FT Domain 439..451  
 FT /note= "putative transmembrane domain"  
 PN WO9745536-A1.  
 PD 04-DEC-1997.  
 PR 30-MAY-1997; U09160.  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA (INDV ) UNIV INDIANA FOUND.  
 PI Bard M, Sturley SP, Yang H;  
 DR WPI: 98-032644/03.  
 DR N-PSDB: T96370.  
 PT Yeast acyl-coenzyme A:cholesterol acyltransferase related enzyme I  
 PT and II - useful to identify inhibitor for treatment of  
 PT hyperlipidaemia, arteriosclerosis and fungal invasion  
 PS Example 1; Page 71-73; 111pp; English.  
 CC Yeast acyl-coenzyme A:cholesterol acyltransferase 2 (ARE2) is

CC encoded by a gene (see T96370) located on chromosome X or XIV that  
 CC was identified by homology to the human acyl-coenzyme A:cholesterol  
 CC acyltransferase I gene (see T96368). ARE2 and ARE1 (see W38417)  
 CC exhibit significant structural and functional homology to the  
 CC putative catalytic component of cholesterol esterification in  
 CC human macrophages. A consensus sequence of the yeast and human  
 CC proteins is given in W38419. Deletion of the ARE1 and ARE2 genes  
 CC produces a viable yeast cell with undetectable esterified sterol.  
 CC A claimed expression vector comprises a nucleic acid molecule  
 CC encoding yeast wild type ARE1 or ARE2 operatively linked to a  
 CC promoter of RNA transcription. The vector and a claimed  
 CC host-vector system can be used for the recombinant production of  
 CC ARE1 or ARE2. The products can be used to identify novel ARE1 or  
 CC ARE2 inhibitors. ARE2 inhibitors can be used to treat  
 CC arteriosclerosis or hyperlipidaemia. ARE1 and ARE2 inhibitors can  
 CC also be used to inhibit fungal growth or to treat fungal invasion  
 CC (all claimed).  
 SQ Sequence 642 AA;

Query Match 31.2%; Score 59; DB 28; Length 642;  
 Best Local Similarity 66.7%; Pred. No. 7.30e+01;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 372 kvigalekscef 383  
 ||| :|||  
 QY 9 KVTGLYERKGEF 20

## RESULT 11

ID W43411 standard; Protein: 642 AA.  
 AC W43411;  
 DT 08-JUN-1998 (first entry)  
 DE Yeast acyl-coenzyme A:cholesterol acyltransferase-related enzyme II.  
 KW Acyl-coenzyme A:cholesterol acyltransferase; ACAT II;  
 KW ACAT related gene product 1; ARGP-1; ARP-2; sterol esterification;  
 KW inhibitor; arteriosclerosis; hyperlipidaemia; yeast.  
 OS Saccharomyces cerevisiae.  
 FH Key Location/Qualifiers  
 FT Domain 215..231  
 FT /note= "transmembrane domain"  
 FT Domain 439..451  
 FT /note= "transmembrane domain"  
 PN WO9745439-A1.  
 PD 04-DEC-1997.  
 PR 30-MAY-1997; U09460.  
 PR 30-MAY-1996; US-657620.  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 PI Sturley SL;  
 DR WPI: 98-032573/03.  
 DR N-PSDB: V01538.  
 PT DNA encoding acyl-coenzyme A: cholesterol acyltransferase II or  
 PT III - useful to identify inhibitors for treatment of  
 PT arteriosclerosis or hyperlipidaemia  
 PS Disclosure; Fig 5C 1-3; 121pp; English.  
 CC This protein comprises yeast acyl-coenzyme A:cholesterol  
 CC acyltransferase-related enzyme II. The invention relates to  
 CC isolated nucleic acids (see V01533-35) coding for human and mouse  
 CC acyl-coenzyme A:cholesterol acyltransferase II and III (see  
 CC W43406-08), also designated ACAT related gene products (ARGP) 1  
 CC and 2. These can be used to identify inhibitors useful in the  
 CC treatment of arteriosclerosis and hyperlipidaemia.  
 SQ Sequence 642 AA;

Query Match 31.2%; Score 59; DB 29; Length 642;  
 Best Local Similarity 66.7%; Pred. No. 7.30e+01;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 372 kvigalekscef 383  
 ||| :|||  
 QY 9 KVTGLYERKGEF 20

## RESULT 12



ID R91391 standard; Peptide; 28 AA.  
AC R91391;  
DT 10-OCT-1996 (first entry)  
DE ME#1 library derived peptide, F15-A9b.  
KW C1s-element; ligand; highly specific DNA binding domain; HSDB; human;  
KW H2-kappa-B; regulatory domain; murine; homologue; IL-6; interleukin;  
KW HLA class I gene; HLA-B27; HLA-J; HLA-A2; acute phase protein response;  
KW syngene; immune response; IL-8; IL-16; diagnosis; gene therapy;  
KW transcription.  
OS Synthetic.  
PN WO9606188-A1.  
PD 29-FEB-1996.  
PF 17-AUG-1995; U10523.  
PR 18-AUG-1994; US-292902.  
PR 15-AUG-1995; US-515190.  
PA (CYTO-) CYTOGEN CORP.  
PI Fowlkes DM, Rodwell JD;  
DR WPI: 96-151391/15.  
PT Synthetic random nucleotide sequences encoding ligand binding  
PT domains - identified by screening library of vectors or peptide(s)  
PT and useful for gene therapy and diagnosis  
PS Claim 70: Page 196: 224pp; English.  
CC The sequences given in R91378-414 represent peptides that bind to  
CC highly specific DNA binding domains (HSDB)'s. These sequences were  
CC tested for binding to the H2-kappa-B oligonucleotide which contains  
CC the NF-kappa-B binding site, and comprises the sequence given in T13579.  
CC These peptides may be used in a composition for diagnosis and gene  
CC therapy and for modifying the transcription or activity of a gene.  
SQ Sequence 28 AA;

Query Match 30.7%; Score 58; DB 18; Length 28;  
Best Local Similarity 31.3%; Pred. No. 9.18e+01;  
Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 10 sfdratgcfhpgckyi 25  
:|::||: ||::

QY 6 TFNKVTGLYKGEFV 21

RESULT 13  
ID P70139 standard; protein; 255 AA.  
AC P70139;  
DT 09-APR-1991 (first entry)  
DE Sequence of a region of the S1 protein of infectious  
DE bronchitis virus (IBV) strain D274.  
KW Poultry vaccine; antigen; immunogen.  
OS Infectious bronchitis virus.  
PN EP-221609-A.  
PD 13-MAY-1987.  
PF 28-OCT-1986; 201884.  
PR 31-OCT-1985; NL-002975.  
PA (DUIN) DUPHAR INT RES BV.  
PI Niesters HGM, Van der Zeijst BAM;.  
DR WPI: 87-130709/19.  
PT New gene coding for protein of infectious bronchitis viruses -  
PT useful in prodn. of peptide or protein for vaccination of poultry  
PT against the viruses.  
PS Claim 2; Fig 2; 21pp; English.  
CC Proteins encoded by IBV strains M41, D1466, D207 or D274 are claimed  
CC together with the genes that encode them, and the gene that encodes  
CC a protein from IBV strain H120. In addition, a peptide is claimed,  
CC characterised in that it comprises at least one of the regions in  
CC which the AA sequences of the IBV strains M41 and M42 are  
CC different, esp. the region between AAs 55 and 70 and/or between AAs  
CC 116 and 131 and/or between AAs 290 and 301.  
SQ Sequence 255 AA;

Query Match 30.7%; Score 58; DB 2; Length 255;  
Best Local Similarity 53.8%; Pred. No. 9.18e+01;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 41 gsyhkcacnfrles 53  
||:|:|:|:|:|

QY 12 GLYKGEFVQLQS 24

RESULT 14  
ID W71512 standard; Protein; 306 AA.  
AC W71512;  
DT 09-NOV-1998 (first entry)  
DE Helicobacter polypeptide GHPO 1346.  
KW GHPO 1346; infection; therapy; diagnosis; vaccine; gastritis;  
KW ulcer.  
OS Helicobacter pylori.  
PN WO9821225-A1.  
PD 22-MAY-1998.  
PF 14-NOV-1997; U21353.  
PR 29-JUL-1997; US-902615.  
PR 14-NOV-1996; US-749051.  
PR 01-APR-1997; US-831309.  
PR 01-APR-1997; US-833457.  
PR 01-APR-1997; US-834705.  
PR 24-JUN-1997; US-881227.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
PI Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C,  
PI Odenbreit S, Tomb J;  
DR WPI: 98-297855/26.  
DR N-PSDB: V52047.  
PT Helicobacter polynucleotide and polypeptide sequences - useful to  
PT treat or prevent gastrointestinal infection  
PS Claim 1; Page 188-189; 362pp; English.  
CC This claimed Helicobacter pylori polypeptide, designated GHPO 1346,  
CC can be used in vaccination methods for preventing or treating  
CC Helicobacter infection. 85 Helicobacter polypeptides (see  
CC W71474-W71558) are claimed, as well as isolated polynucleotides  
CC (see V52009-93) that encode them. The invention also provides:  
CC methods for producing these Helicobacter polypeptides in  
CC recombinant host systems, and related expression cassettes, vectors  
CC and transformed or transfected host cells; live vaccine vectors  
CC that contain the polynucleotides of the invention and which can be  
CC used to prevent or treat Helicobacter infection; therapeutic and/or  
CC prophylactic methods involving administration of polynucleotide  
CC molecules, polypeptides or monospecific antibodies; methods for  
CC detecting the presence of Helicobacter in samples using e.g.  
CC the polypeptides or monospecific antibodies; and methods for  
CC purifying the polypeptides by antibody-based affinity  
CC chromatography.  
SQ Sequence 306 AA;

Query Match 30.7%; Score 58; DB 35; Length 306;  
Best Local Similarity 27.8%; Pred. No. 9.18e+01;  
Matches 5; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Db 40 krgdyhkavafykrcsnl 57  
||:|:|:|:|:|

QY 3 KRCTFNKVTGLYKGEFV 20

RESULT 15  
ID W04216 standard; Protein; 388 AA.  
AC W04216;  
DT 23-NOV-1996 (first entry)  
DE Rat superior cervical ganglion p2x receptor.  
KW ATP P2x receptor; purinoceptor; ligand-gated ion channel; agonist;  
KW antagonist; epilepsy; cognition; emesis; pain; asthma;  
KW peripheral vascular disease; hypertension; irritable bowel syndrome;  
KW premature ejaculation; cystitis; therapy.  
OS Rattus sp.  
PN WO9533048-A2.  
PD 07-DEC-1995.  
PF 24-MAY-1995; E01968.  
PR 27-MAY-1994; GB-010664.  
PR 09-FEB-1995; GB-002480.  
PA (GLAX) GLAXO GROUP LTD.

Search completed: Mon Aug 23 13:07:11 1999  
Job time : 26 secs.

\*\*\*\*\*

```

KEYWORDS      DNA biosynthesis; nucleotidyltransferase
SUMMARY       #length 788 #molecular-weight 90070 #checksum 8976

Query Match   40.1%; Score 77; DB 1; Length 788;
Best Local Similarity 61.1%; Pred. No. 1.05e-02;
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Db 86 RHLKGLTGLYQMKGCEF 103
   | : | : | : | : | : |
Qy 4 RCTFNKVTGLYE-KGCEF 20

RESULT 3
ENTRY   B70153 #type complete
TITLE   conserved hypothetical protein BB0427 - Lyme disease
        spirochete
ORGANISM #formal_name Borrelia burgdorferi #common_name Lyme disease
        spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
        13-Sep-1998
ACCESSION B70153
REFERENCE A70100
AUTHORS Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
        Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
        Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
        J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
        Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
        Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
        J.; Utterback, T.; Watthey, L.; McDonald, L.; Artach, P.;
        Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
        K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
        Nature (1997) 390:580-586
        Genomic sequence of a Lyme disease spirochaete, Borrelia
        burgdorferi.
#cross-references NCBI:98065943
#accession B70153
#status preliminary; nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-263 #label KLF
#cross-references GB:AE001148; GB:AE000783; NID:g2688336; PID:g2688342;
        TIGR:BB0427
#experimental_source strain B31
CLASSIFICATION #superfamily conserved hypothetical protein MG056
SUMMARY #length 263 #molecular-weight 29739 #checksum 6568

Query Match 37.0%; Score 70; DB 2; Length 263;
Best Local Similarity 43.8%; Pred. No. 2.11e-01;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 180 FKRTAEYLKRGDAFVL 195
   | : : : | : | : |
Qy 7 FKNVTGLYKRGCEFVL 22

RESULT 4
ENTRY   C64175 #type complete
TITLE   hypothetical protein HI1895 (lsg locus) - Haemophilus
        influenzae
ORGANISM #formal_name Haemophilus influenzae
        #18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
        17-Oct-1997
ACCESSION C64175; S27582
REFERENCE A64000
AUTHORS Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
        Kirkness, E.F.; Kerlavage, A.R.; Bult, C.J.; Tomb, J.F.;
        Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
        FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
        Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
        J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
        M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.T.; Saudek,
        D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
        J.L.; Geodhagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.;

```

TITLE	titin - rabbit (fragment)	
ORGANISM	#formal_name Oryctolagus cuniculus #common_name domestic rabbit	
DATE	14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 12-Sep-1997	
ACCESSIONS	I4521	
REFERENCE	I4520	
#authors	Labeit, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Hsieh, C.L.; Francke, U.; Leonard, K.; Wardale, J.; Whiting, A.; Trinick, J.	
#journal	Nature (1990) 345:273-276	
#title	A regular pattern of two types of 100-residue motif in the sequence of titin.	
#cross-references	MUID:90238553	
#accession	I4521	
#status	preliminary: translated from GB/EMBL/DBDJ	
#molecule_type	mRNA	
#residues	1-1000 #label LAB	
#cross-references	EMBL:X17330; NID:gl758; PID:g930252	
CLASSIFICATION	#superfamily titin; fibronectin type III repeat homology; immunoglobulin homology; protein kinase homology	
SUMMARY	#length 1000 #checksum 1229	
Query Match	36.08; Score 68; DB 2; Length 1000;	
Best Local Similarity	64.78; Pred. No. 4.82e-01;	
Matches	11; Conservative 3; Mismatches 1; Indels 2; Gaps 2;	
Db	854 KTY-KVTGLSE-GCEY 868	
QY	4 RCIFNKVTGLYKGEF 20	
	:::        :::	
RESULT	7	
ENTRY	S20901 #type fragment	
TITLE	titin - rabbit (fragment)	
ORGANISM	#formal_name Oryctolagus cuniculus #common_name domestic rabbit	
DATE	30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 12-Sep-1997	
ACCESSIONS	S20901: I46520	
REFERENCE	S20897	
#authors	Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.	
#journal	EMBO J. (1992) 11:1711-1716	
#title	Towards a molecular understanding of titin.	
#cross-references	MUID:92258380	
#accession	S20901	
#status	nucleic acid sequence not shown; translation not shown	
#molecule_type	mRNA	
#residues	1-6805 #label LAB	
#cross-references	EMBL:X64696	
##note	the nucleotide sequence was submitted to the EMBL Data Library, February 1992	
REFERENCE	I46520	
#authors	Labeit, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Hsieh, C.L.; Francke, U.; Leonard, K.; Wardale, J.; Whiting, A.; Trinick, J.	
#journal	Nature (1990) 345:273-276	
#title	A regular pattern of two types of 100-residue motif in the sequence of titin.	
#cross-references	MUID:90238553	
#accession	I46520	
#status	translated from GB/EMBL/DBDJ	
#molecule_type	mRNA	
#residues	4235-5250 #label LA2	
#cross-references	EMBL:X17329; NID:gl756; PID:g930251	
CLASSIFICATION	#superfamily titin; fibronectin type III repeat homology; immunoglobulin homology; protein kinase homology	
KEYWORDS	muscle	
SUMMARY	#length 6805 #checksum 1434	
Query Match	36.08; Score 68; DB 2; Length 6805;	
Best Local Similarity	64.78; Pred. No. 4.82e-01;	
Matches	11; Conservative 3; Mismatches 1; Indels 2; Gaps 2;	

```

##cross-references EMBL:X92412; NID:g1236761
REFERENCE
S37393
#authors
Gautel, M.; Leonard, K.; Labeit, S.
#journal
EMBO J. (1993) 12:3827-3834
#title
Phosphorylation of KSP motifs in the C-terminal region of
titin in differentiating myoblasts.
#cross-references MUID:94008990
#accession
S37393
##molecule_type mRNA
##residues
26831-26926 ##label GAU
REFERENCE
A66736
#authors
Improra, S.; Politou, A.S.; Pastore, A.
#submission
submitted to the Brookhaven Protein Data Bank, February 1996
#cross-references PDB:1TIT
#contents
annotation: conformation by (1)H-NMR, residues 5253-5341
REFERENCE
A66201
#authors
Pfuhl, M.; Pastore, A.
#submission
submitted to the Brookhaven Protein Data Bank, August 1996
#cross-references PDB:1NCT
#contents
annotation: conformation by (1)H-NMR, residues 'S',
26059-26155
GENETICS
#gene
GDB:TTN
##cross-references GDB:127867; OMIM:188840
#map_position 2q31-2q32
FUNCTION
structural protein forming filaments in striated muscle
#description
#superfamily titin; fibronectin type III repeat homology;
immunoglobulin homology; protein kinase homology
CLASSIFICATION
alternative splicing; calmodulin binding; cardiac muscle;
duplication; glycoprotein; phosphoprotein;
phosphotransferase; serine/threonine-specific protein
kinase; skeletal muscle; structural protein
KEYWORDS
FEATURE
24752-25008
#domain protein kinase homology #label KIN
84,177,905,2276,
2378,2459,2481,
2563,2669,2763,
2896,3088,3179,
3384,3432,3628,
3772,4068,4318,
5047,5246,5823,
6213,6264,6899,
6800,7083,7300,
7506,7597,8338,
8447,8455,8719,
8938,9375,10130,
10133,10177,10398,
11066,11488,11515,
11635,11949,12170,
12478,12526,12645,
12875,13001,13036,
13295,13540,13787,
13986,14085,14529,
14870,14910,15458,
15704,16397,16520,
16695
#binding_site carbohydrate (Asn) (covalent) #status
predicted\
16780,16976,17579,
17602,17667,17681,
17845,17899,18121,
18188,18209,18336,
18670,18680,18685,
18707,18945,19198,
19238,19346,19753,
19786,20034,20056,
20183,20324,20426,
20840,21270,21407,
21538,21566,21665,
21900,21935,22295,
22495,22627,22897,
23024,23318,23883,
24012,24177,24290,

```

```

24447,24642,24840,
25201,25927,26133,
26283,26682,26698,
26811
#binding_site carbohydrate (Asn) (covalent) #status
predicted\
26171,26178,26184,
26190
#binding_site phosphate (Ser) (covalent) #status
experimental
SUMMARY
#length 26926 #molecular-weight 2993498 #checksum 431
Query Match 36.0%; Score 68; DB 1; Length 26926;
Best Local Similarity 64.7%; Pred. No. 4,82e-01;
Matches 11; Conservative 3; Mismatches 1; Indels 2; Gaps 2;
Oy 4 RCTFNKVTGLYEKGCEP 20
Db 14426 KCTY-KVTGLSE-GCEY 14440
:|||||:
:|||||:
Oy 4 RCTFNKVTGLYEKGCEP 20
RESULT 9
ENTRY H70761 #type complete
TITLE probable dnapolymerase III - Mycobacterium tuberculosis
(strain H37Rv)
ORGANISM #formal_name Mycobacterium tuberculosis
17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
17-Jul-1998
ACCESSIONS H70761
REFERENCE A70500
#authors
Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry
III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrell, B.G.
#journal
Nature (1998) 393:537-544
#title
Deciphering the biology of Mycobacterium tuberculosis from
the complete genome sequence.
#cross-references MUID:98295987
#accession H70761
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-1184 #label COL
##cross-references GB:274020; GB:AL123456; NID:g3261584; PID:e248547;
#experimental_source strain H37Rv
GENETICS
#gene
dnaEI
SUMMARY
#length 1184 #molecular-weight 129322 #checksum 2714
Query Match 35.4%; Score 67; DB 2; Length 1184;
Best Local Similarity 58.3%; Pred. No. 7,24e-01;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Oy 2 TKRCTFNKVTGL 13
Db 1118 TRCFTDKVSAL 1129
:|||||:
:|||||:
Oy 2 TKRCTFNKVTGL 13
RESULT 10
ENTRY G69366 #type complete
TITLE homoserine dehydrogenase (hom) homolog - Archaeoglobus
fulgidus
ORGANISM #formal_name Archaeoglobus fulgidus
05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
21-Aug-1998
ACCESSIONS G69366
REFERENCE A69250
#authors
Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,

```

K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kervagave, A.R.; Graham, D.E.; Kyripides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodok, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.

Nature (1997) 390:364-370

The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*.

#cross-references MUID:98049343

#accession G69366

#status preliminary; nucleic acid sequence not shown; translation not shown

##molecule\_type DNA

##residues 1-326 #label KLF

##cross-references GB:AE001039; GB:AE000782; NID:g2689362; PID:g2649661; TIGR:AF0935

CLASSIFICATION #superfamily homoserine dehydrogenase; homoserine dehydrogenase homology

SUMMARY #length 326 #molecular-weight 35431 #checksum 626

Query Match 34.9%; Score 66; DB 2; Length 326;

Best Local Similarity 27.3%; Pred. No. 1.08e+00;

Matches 6; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Db 155 YLALCEESVKGIFNGTCNVIL 176

QY 1 YTRCTFNKVTGLYKGCFFVL 22

RESULT 11

ENTRY S58652 #type complete

TITLE hypothetical protein YFR036w-a - yeast (*Saccharomyces cerevisiae*)

ORGANISM #formal\_name *Saccharomyces cerevisiae*

DATE 27-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 05-Dec-1997

ACCESSIONS S58652

REFERENCE S58186

#authors Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiya, Y.; Soeda, E.; Yokoyama, K.; Yamazaki, M.; Tashiro, H.; Eki, T.

#submission submitted to the EMBL Data Library, May 1995

#description Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*.

#accession S58652

##molecule\_type DNA

##residues 1-216 #label MUR

##cross-references EMBL:D50617; MIPS:YFR036w-a

GENETICS

#map-position 6R

SUMMARY #length 216 #molecular-weight 24214 #checksum 6341

Query Match 34.4%; Score 65; DB 2; Length 216;

Best Local Similarity 44.4%; Pred. No. 1.62e+00;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 21 FYNDHGLYTWGCFNAMES 38

QY 7 FNKVTGLYKGCFFVLQS 24

RESULT 12

ENTRY T00100 #type complete

TITLE hypothetical protein 4 - *Actinobacillus actinomycetemcomitans*

ORGANISM #formal\_name *Actinobacillus actinomycetemcomitans*

DATE 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change

K.E.; Ketchum, K.A.; Dodson, R.J.; Gwinn, M.; Hickey, E.K.; Peterson, J.D.; Richardson, D.L.; Kervagave, A.R.; Graham, D.E.; Kyripides, N.C.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Dougherty, B.A.; McKenny, K.; Adams, M.D.; Loftus, B.; Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.; Glodok, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.; D'Andrea, K.P.; Bowman, C.; Fujii, C.; Garland, S.A.; Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.

Nature (1997) 390:364-370

The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Archaeoglobus fulgidus*.

#cross-references MUID:98049343

#accession G69366

#status preliminary; nucleic acid sequence not shown; translation not shown

##molecule\_type DNA

##residues 1-326 #label KLF

##cross-references GB:AE001039; GB:AE000782; NID:g2689362; PID:g2649661; TIGR:AF0935

CLASSIFICATION #superfamily homoserine dehydrogenase; homoserine dehydrogenase homology

SUMMARY #length 326 #molecular-weight 35431 #checksum 626

Query Match 34.9%; Score 66; DB 2; Length 326;

Best Local Similarity 27.3%; Pred. No. 1.08e+00;

Matches 6; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Db 155 YLALCEESVKGIFNGTCNVIL 176

QY 1 YTRCTFNKVTGLYKGCFFVL 22

RESULT 11

ENTRY S58652 #type complete

TITLE hypothetical protein YFR036w-a - yeast (*Saccharomyces cerevisiae*)

ORGANISM #formal\_name *Saccharomyces cerevisiae*

DATE 27-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 05-Dec-1997

ACCESSIONS S58652

REFERENCE S58186

#authors Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiya, Y.; Soeda, E.; Yokoyama, K.; Yamazaki, M.; Tashiro, H.; Eki, T.

#submission submitted to the EMBL Data Library, May 1995

#description Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces cerevisiae*.

#accession S58652

##molecule\_type DNA

##residues 1-216 #label MUR

##cross-references EMBL:D50617; MIPS:YFR036w-a

GENETICS

#map-position 6R

SUMMARY #length 216 #molecular-weight 24214 #checksum 6341

Query Match 34.4%; Score 65; DB 2; Length 216;

Best Local Similarity 44.4%; Pred. No. 1.62e+00;

Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 21 FYNDHGLYTWGCFNAMES 38

QY 7 FNKVTGLYKGCFFVLQS 24

RESULT 12

ENTRY T00100 #type complete

TITLE hypothetical protein 4 - *Actinobacillus actinomycetemcomitans*

ORGANISM #formal\_name *Actinobacillus actinomycetemcomitans*

DATE 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change

22-Jan-1999

ACCESSIONS T00100

REFERENCE Z14111

#authors Nakano, Y.; Yoshida, Y.; Yamashita, Y.; Koga, T.

#journal Biochim. Biophys. Acta (1998) 1442:409-414

#title A gene cluster for 6-deoxy-L-talan synthesis in *Actinobacillus actinomycetemcomitans*.

#accession T00100

#status preliminary; translated from GB/EMBL/DBJ

##molecule\_type DNA

##residues 1-261 #label NAK

##cross-references EMBL:AB010415; NID:dl194915; PID:dl029069

##experimental\_source strain NCTC9710

SUMMARY #length 261 #molecular-weight 29525 #checksum 4957

Query Match 34.4%; Score 65; DB 2; Length 261;

Best Local Similarity 53.3%; Pred. No. 1.62e+00;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 144 KRCFENHMTVGYSKS 158

QY 3 KRCTFNKVTGLYERK 17

RESULT 13

ENTRY E70455 #type complete

TITLE sulfur oxidation protein SoxB - Aquifex aeolicus

ORGANISM #formal\_name *Aquifex aeolicus*

DATE 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 08-May-1998

ACCESSIONS E70455

REFERENCE A70300

#authors Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.; Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short, J.M.; Olson, G.J.; Swanson, R.V.

#journal Nature (1998) 392:353-358

#title The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

#cross-references MUID:98196666

#accession E70455

#status preliminary; nucleic acid sequence not shown; translation not shown

##molecule\_type DNA

##residues 1-592 #label AQF

##cross-references GB:AE000757; NID:g2984092; PID:g2984096; GB:AE000657

GENETICS

#gene soxB

SUMMARY #length 592 #molecular-weight 66379 #checksum 4040

Query Match 34.4%; Score 65; DB 2; Length 592;

Best Local Similarity 33.3%; Pred. No. 1.62e+00;

Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Db 181 FGKDIVLKRIKELEBAGCEPISON 204

QY 1 YTRCTFNKVTGLYKGCFFVLQS 24

RESULT 14

ENTRY TRT1 #type complete

TITLE trypsin (EC 3.4.21.4) I precursor - rat

ALTERNATE\_NAMES trypsinogen I

ORGANISM #formal\_name *Rattus norvegicus* #common\_name Norway rat

DATE 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 20-Mar-1998

ACCESSIONS B22657; A00948

REFERENCE A22657

#authors Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.

#journal J. Biol. Chem. (1984) 259:14255-14264

#title Structure of two related rat pancreatic trypsin genes.

```
#genome      plasmid pXOI
KEYWORDS     DNA binding; isomerase
SUMMARY      #length 887 #molecular-weight 101963 #checksum 7198

Query Match          33.9%   Score 64; DB 2; Length 887;
Best Local Similarity 45.5%; Pred.No. 2.41e+00;
Matches              5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```



\*\*\*\*\*

W A R E H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mpsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:04:59 1999; MasPar time 5.10 Seconds  
Tabular output not generated. 133.146 Million cell updates/sec

Title: >US-09-049-696-43  
Description: (1-24) from US09049696.pep  
Perfect Score: 189  
Sequence: 1 YTKRCTFNKVTGLYKGCCEFLQS 24

Scoring table: PAM 150  
Gap 15

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot37  
1:swissprot

Statistics: Mean 29.651; Variance 38.303; scale 0.774

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	77	40.7	788	1	DPOL_HPBHE DNA POLYMERASE (EC 2.7	1.53e-03
2	69	36.5	903	1	ECLC_BOVIN EPITHELIAL CHLORIDE CH	7.18e-02
3	67	35.4	1184	1	DP3A_MYCTU PROBABLE DNA POLYMERAS	1.80e-01
4	65	34.4	292	1	TFZ_HUMAN TAFAZIN.	4.45e-01
5	64	33.9	246	1	TRY1_RAT TRYPsinogen I, ANIONIC	6.94e-01
6	64	33.9	887	1	TOPI_BACAN DNA TOPOISOMERASE I (E	6.94e-01
7	63	33.3	162	1	AROK_HELPY SHIKIMATE KINASE (EC 2	1.08e+00
8	63	33.3	788	1	DPOL_HPBHW DNA POLYMERASE (EC 2.7	1.08e+00
9	63	33.3	836	1	DPOL_HPBHU DNA POLYMERASE (EC 2.7	1.08e+00
10	62	32.8	787	1	DPOL_HPBDC DNA POLYMERASE (EC 2.7	1.66e+00
11	62	32.8	788	1	DPOL_HPBDB DNA POLYMERASE (EC 2.7	1.66e+00
12	61	32.3	199	1	RSBX_BACSU SIGMA-B NEGATIVE EFEC	2.56e+00
13	61	32.3	296	1	ULOT_HSVII PROTEIN UL7.	2.56e+00
14	61	32.3	313	1	YA71_METJA HYPOTHETICAL PROTEIN M	2.56e+00
15	61	32.3	405	1	SAPB_HUMAN SERUM RESPONSE FACTOR	2.56e+00
16	61	32.3	2787	1	TELL_YEAST TELOMER LENGTH REGULAT	2.56e+00
17	60	31.7	150	1	DEF_CIOAB POLYPEPTIDE DEFORMYLAS	3.91e+00
18	60	31.7	246	1	TRY2_RAT TRYPsinogen II, ANIONI	3.91e+00
19	60	31.7	247	1	TRY2_CANFA TRYPsinogen, ANIONIC P	3.91e+00
20	60	31.7	375	1	PCE_TACTR PROCTOTTING ENZYME PRE	3.91e+00
21	60	31.7	432	1	SYWC_YEAST TRYPTOPHANYL-TRNA SYNT	3.91e+00
22	60	31.7	477	1	NIFD_METMP NITROGENASE MOLYBDENUM	3.91e+00
23	60	31.7	547	1	GTRI_LEIDO MEMBRANE TRANSPORTER D	3.91e+00

24	59	31.2	248	1	TRY3_CHICK TRYPsinogen II-P29 PRE	5.94e+00
25	59	31.2	306	1	SPEB_ECOLI AGMATINASE (EC 3.5.3.1	5.94e+00
26	59	31.2	336	1	Y883_METJA HYPOTHETICAL PROTEIN M	5.94e+00
27	59	31.2	529	1	ZACE_HUMAN PROTEIN PHOSPHATASE PP	5.94e+00
28	59	31.2	642	1	STEROL_O-ACYLTRANSFERA	5.94e+00
29	59	31.2	1054	1	TREA_YEAST ACID TREHALASE PRECURS	5.94e+00
30	59	31.2	2852	1	RRPB_IBVB RNA-DIRECTED RNA POLYM	5.94e+00
31	58	30.7	138	1	PA2B_TRIFL PHOSPHOLIPASE 'A2, BASI	8.98e+00
32	58	30.7	244	1	TRY2_XENLA TRYPsinogen PRECURSOR.	8.98e+00
33	58	30.7	306	1	BLAC_STRFR BETA-LACTAMASE PRECURS	8.98e+00
34	58	30.7	388	1	P2X4_RAT P2X PURINOCEPTOR 4 (AT	8.98e+00
35	58	30.7	425	1	YGGP_ECOLI HYPOTHETICAL 45.8 KD P	8.98e+00
36	58	30.7	462	1	DRTS_PARTI DIHYDROFOLATE REDUCTAS	8.98e+00
37	58	30.7	855	1	ENV_FIVU8 ENV POLYPROTEIN PRECUR	8.98e+00
38	58	30.7	923	1	YOL3_CABEL HYPOTHETICAL 103.9 KD	8.98e+00
39	58	30.7	1067	1	IMB4_SCHPO PUTATIVE IMPORTIN BETA	8.98e+00
40	57	30.2	286	1	DMA_HAEIN DNA ADENINE METHYLASE	1.35e+01
41	57	30.2	388	1	P2X4_HUMAN P2X PURINOCEPTOR 4 (AT	1.35e+01
42	57	30.2	470	1	BFRI_YEAST NUCLEAR SEGREGATION PR	1.35e+01
43	57	30.2	1095	1	PIPA_DROME 1-PHOSPHATIDYLINOSITOL	1.35e+01
44	57	30.2	1191	1	DNBI_MCMVS MAJOR DNA-BINDING PROT	1.35e+01
45	57	30.2	1202	1	YE01_SCHPO HYPOTHETICAL 138.5 KD	1.35e+01

ALIGNMENTS

RESULT 1  
ID DPOL\_HPBHE STANDARD; PRT; 788 AA.  
AC P13846;  
DT 01-JAN-1990 (REL. 13, CREATED)  
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.  
OS HERON HEPATITIS B VIRUS.  
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 88333160.  
RA SPRENGEL R., KALETA E.F., WILL H.;  
RT "Isolation and characterization of a hepatitis B virus endemic in herons."  
RL J. VIROL. 62:3832-3839(1988).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
EMBL; M22056; G325454; -  
DR PIR; A30082; JVDLHH.  
DR PFAM; PF00078; rvt; 1.  
DR PFAM; PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM; PF00336; DNA\_pol\_viral\_C; 1.  
SQ TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.  
KW SEQUENCE 788 AA; 90070 MW; CC3BC3AD CRC32;  
Query Match 40.7%; Score 77; DB 1; Length 788;  
Best Local Similarity 61.1%; Pred. No. 1.53e-03;  
Matches 11; Conservative 3; Mismatches 3; Indels 1; Gaps 1;  
Db 86 RHKLGKLTGLYQMKGEF 103  
QY 4 RCTFNKVTGLYE-KGCEF 20  
RESULT 2  
ID ECLC\_BOVIN STANDARD; PRT; 903 AA.  
AC P54281;  
DT 01-OCT-1996 (REL. 34, CREATED)

```

RC STRAIN=H37RV;
RA MURPHY L., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC =====
CC EMBL: Z74020; E248547; .
CC TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.
CC SEQUENCE 1184 AA; 129322 MW; D7EFB285 CRC32;
CC
CC Query Match 35.48; Score 67; DB 1; Length 1184;
CC Best Local Similarity 58.38; Pred. No. 1.80e-01;
CC Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
CC
CC Db 1118 TROCTFDKVSAL 1129
CC QY 2 TKRCTFNKVTGL 13
CC
CC RESULT 4
CC ID TFZ_HUMAN STANDARD; PRT; 292 AA.
CC AC Q16635;
CC DT 01-NOV-1997 (REL. 35, CREATED)
CC DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
CC DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
CC DE TAFAZZIN.
CC GN TAZ OR BFE2 OR G4.5.
CC OS HOMO SAPIENS (HUMAN).
CC OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
CC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
CC [1]
CC RN SEQUENCE FROM N.A.
CC RC TISSUE-SKELETAL MUSCLE, AND HEART;
CC RX MEDLINE; 96224398.
CC RA BIONE S., D'ADAMO P., MAESTRINI E., GEDEON A.K., BOLHUIS P.A.,
CC TONIOLO D.;
CC RA "A novel X-linked gene, G4.5, is responsible for Barth syndrome.";
CC RL NAT. GENET. 12:385-399(1996).
CC CC -1- SUBCELLULAR LOCATION: ISOFORMS WITH HYDROPHOBIC N-TERMINUS ARE
CC THOUGHT TO BE MEMBRANE ANCHORED. SHORTEST FORMS, LACKING THE
CC HYDROPHOBIC STRETCH, MAY BE SOLUBLE CYTOPLASMIC PROTEINS.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN CARDIAC AND SKELETAL MUSCLE. UP
CC TO 10 ISOFORMS CAN BE PRESENT IN DIFFERENT AMOUNTS IN DIFFERENT
CC TISSUES. MOST ISOFORMS ARE UBQUITOUS. ISOFORMS THAT LACK THE N-
CC TERMINUS ARE FOUND IN LEUKOCYTES AND FIBROBLASTS, BUT NOT IN HEART
CC AND SKELETAL MUSCLE. SOME FORMS APPEAR RESTRICTED TO CARDIAC AND
CC SKELETAL MUSCLE OR TO LEUKOCYTES.
CC -1- DOMAIN: THE HYDROPHILIC DOMAIN MAY SERVE AS AN EXPOSED LOOP
CC INTERACTING WITH OTHER PROTEINS.
CC -1- DISEASE: DEFECTS IN TAZ ARE THE CAUSE OF BARTH SYNDROME (BTHS), A
CC IS A SEVERE INHERITED DISORDER, OFTEN FATAL IN CHILDHOOD,
CC CHARACTERIZED BY CARDIAC AND SKELETAL MYOPATHY, SHORT STATURE AND
CC NEUTROPENIA.
CC -1- ALTERNATIVE PRODUCTS: UP TO 10 ISOFORMS. SHORTER TAFAZZINS START
CC WITH THE SECOND METHIONINE AT POS. 25. ALTERNATIVE SPLICING OF THE
CC HYDROPHILIC CENTRAL REGION (EXONS 5-7) MAY PRODUCE 5 VARIANTS. TWO
CC ISOFORMS, CONTAINING ALL THREE EXONS OR LACKING EXON 5 ONLY ARE
CC CONSISTANTLY ABUNDANT. THE SEQUENCE SHOWN HERE IS THAT OF THE
CC LONGEST ISOFORM.
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: X92763; E209047; -  
DR EMBL: X92764; E209047; JOINED.  
DR EMBL: X92762; E209046; -  
DR MIM: 302060; -  
DR MIM: 300069; -  
KW ALTERNATIVE SPLICING; TRANSMEMBRANE.  
FT TRANSMEM 1 30 MEMBRANE ANCHORED (POTENTIAL).  
FT DOMAIN 124 194 HYDROPHILIC.  
FT VARSPLIC 124 153 MISSING (IN ISOFORMS WITHOUT EXON 5).  
FT VARSPLIC 154 180 MISSING (IN ISOFORMS WITHOUT EXON 6).  
FT VARSPLIC 181 194 MISSING (IN ISOFORMS WITHOUT EXON 7).  
SQ SEQUENCE 292 AA; 33459 MW; B6835375 CRC32;

Query Match 34.48; Score 65; DB 1; Length 292;

Best Local Similarity 43.84; Pred. No. 4.49e-01;

Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 152 EKGGVYKGMDFILE 167

QY 8 NKTGLYEKGCERVLQ 23

RESULT 5

ID TRY1\_RAT STANDARD; PRT; 246 AA.

AC P00762;

DT 21-JUL-1986 (REL. 01, CREATED)

DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE TRYPSINOGEN I, ANIONIC PRECURSOR (EC 3.4.21.4) (PRETRYPSINOGEN I).

OS RATTUS NORVEGICUS (RAT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.

[1]

SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RX MEDLINE: 82265624.

RA McDONALD R.J., STARY S.J., SWIFT G.H.;

RT "Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide

sequences of the cloned cDNAs."

RL J. BIOL. CHEM. 257:9724-9732(1982).

[2]

SEQUENCE FROM N.A.

RX MEDLINE: 85054880.

RA CRAIK C.S., CHOO Q.L., SWIFT G.H., QUINTO C., McDONALD R.J.,

RA RUTTER W.J.;

RT "Structure of two related rat pancreatic trypsin genes."

RL J. BIOL. CHEM. 259:14255-14264(1984).

[3]

X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).

RX MEDLINE: 87292123.

RA SPRANG S., STANDING T., FLETTERICK R.J., STROUD R.M., FINER-MOORE J.,

RA XUONG N.-H., HAMLIN R., RUTTER W.J., CRAIK C.S.;

RT "The three-dimensional structure of Asn102 mutant of trypsin: role of

Asp102 in serine protease catalysis."

RL SCIENCE 237:905-909(1987).

CC -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.

CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -!- THIS SEQUENCE REPRESENTS THE PRECURSOR OF THE MAJOR FORM OF

TRYPSIN PRODUCED BY THE ADULT PANCREAS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE

TRYPSIN FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: V01273; G57408; -  
DR EMBL: J00778; G206508; -  
DR PIR: A00948; TRRT1.  
DR PDB: 1TRM; 15-JUL-93.  
DR PDB: 2TRM; 16-JUL-88.  
DR PDB: 1BRA; 30-APR-94.  
DR PDB: 1BRB; 31-JUL-94.  
DR PDB: 1BRC; 31-MAY-94.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
DR PFAM: PF00089; trypsin; 1.  
KW HYDROLASE; SERINE PROTEASE; DIGESTION; PANCREAS; ZIMOGEN; SIGNAL;  
KW MULTIGENE FAMILY; 3D-STRUCTURE.  
FT SIGNAL 1 15  
FT PROPEP 16 23 ACTIVATION PEPTIDE.  
FT CHAIN 24 246 TRYPSIN I, ANIONIC.  
FT ACT\_SITE 63 63 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 107 107 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 200 200 CHARGE RELAY SYSTEM.  
FT DISULFID 30 160  
FT DISULFID 48 64  
FT DISULFID 132 233  
FT DISULFID 139 206  
FT DISULFID 171 185  
FT DISULFID 196 220  
FT SITE 194 194  
FT STRAND 25 25  
FT STRAND 28 29  
FT TURN 32 33  
FT TURN 36 37  
FT STRAND 38 42  
FT STRAND 46 54  
FT TURN 55 56  
FT STRAND 57 60  
FT HELIX 62 64  
FT STRAND 70 73  
FT STRAND 77 77  
FT TURN 78 79  
FT STRAND 86 95  
FT TURN 97 98  
FT STRAND 100 100  
FT TURN 101 104  
FT STRAND 105 105  
FT TURN 106 106  
FT STRAND 109 113  
FT STRAND 127 127  
FT TURN 135 136  
FT STRAND 138 143  
FT STRAND 157 157  
FT STRAND 159 165  
FT STRAND 168 174  
FT TURN 176 178  
FT TURN 181 182  
FT STRAND 183 186  
FT TURN 189 190  
FT STRAND 194 194  
FT TURN 197 198  
FT TURN 200 201  
FT STRAND 203 206  
FT TURN 207 208  
FT STRAND 209 216  
FT TURN 223 224  
FT STRAND 227 231  
FT HELIX 232 234  
FT HELIX 236 244  
FT TURN 245 245  
SQ SEQUENCE 246 AA; 25959 MW; C6D097E0 CRC32;

Query Match 33.9%; Score 64; DB 1; Length 246;  
Best Local Similarity 41.2%; Pred. No. 6.94e-01;  
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Db 220 CALPONPGVYTKVCNFV 236  
QY 5 CIFNKVTGLYKGEFV 21

RESULT 6  
ID TOPIBACAN STANDARD; PRT: 887 AA.  
AC P40114;  
DT 01-FEB-1995 (REL. 31, CREATED)  
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE DNA TOPOISOMERASE I (EC 5.99.1.2) (OMEGA-PROTEIN) (RELAXING ENZYME)  
DE (UNWISTING ENZYME) (SWIVELASE).  
GN TOPX.  
OS BACILLUS ANTHRACIS.  
OG PLASMID PXO1.  
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;  
OC BACILLUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=7702;  
RX MEDLINE; 94203050.  
RA FOUET A., SIRARD J.-C., MOCK M.;  
RT "Bacillus anthracis pxo1 virulence plasmid encodes a type 1 DNA  
topoisomerase.";  
RL MOL. MICROBIOL. 11:471-479(1994).  
CC -!- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE  
CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER. MAY PLAY  
AN IMPORTANT ROLE IN EITHER THE MAINTENANCE OF, OR EXPRESSION  
FROM, THE VIRULENCE PLASMID. ITS INFLUENCE ON EXPRESSION COULD BE  
SPECIFIC. TUNING LOCAL SUPERHELICITY OR, MORE GENERALLY, MODIFYING  
THE SUPERHELICITY OF THE ENTIRE PLASMID IN RESPONSE TO  
ENVIRONMENTAL STIMULI, THUS COULD BE A FUNDAMENTAL FACTOR INVOLVED  
IN VIRULENCE.  
CC -!- CATALYTIC ACTIVITY: ATP-INDEPENDENT BREAKAGE OF SINGLE-STRANDED  
DNA, FOLLOWED BY PASSAGE AND REJOINING.  
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -!- WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT  
SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN  
IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE  
ENZYME-SEVERED DNA STRAND.  
CC -!- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE  
FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M97227; G478986; -  
DR PROSITE; PS00396; TOPOISOMERASE\_I\_PROK; 1.  
DR PFAM; PF01131; Topoisom\_bac; 1.  
DR HSP; P06612; 1ECL.  
KW ISOMERASE; TOPOISOMERASE; DNA-BINDING; PLASMID; VIRULENCE;  
ZINC-FINGER.  
FT ACT\_SITE 316 316 DNA CLEAVAGE (BY SIMILARITY).  
FT ZN\_FING 620 644 C4-TYPE (POTENTIAL).  
FT ZN\_FING 710 734 C4-TYPE (POTENTIAL).  
FT ZN\_FING 801 824 C4-TYPE (POTENTIAL).  
SQ SEQUENCE 887 AA; 101963 MW; 12F1B3B5 CRC32;

Query Match 33.9%; Score 64; DB 1; Length 887;  
Best Local Similarity 45.5%; Pred. No. 6.94e-01;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 730 YONGCDFMLPA 740

QY 14 YEKGCEFVLIQS 24

RESULT 7  
ID AROK\_HELPY STANDARD; PRT: 162 AA.  
AC P58073;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE SHIKIMATE KINASE (EC 2.7.1.71) (SK).  
GN AROK OR HP0157.  
OS HELICOBACTER PYLORI (CAMPYLOBACTER PYLORI).  
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; HELICOBACTER GROUP;  
RN HELICOBACTER.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE; 97394467.  
RA TOMB J.-F., WHITE O., KERLAVAGE A.R., CLAYTON R.A., SUTTON G.G.,  
RA FLEISCHMANN R.D., KETCHUM K.A., KLENK H.-P., GILL S., DOUGHERTY B.A.,  
RA NELSON K., QUACKENBUSH J., ZHOU L., KIRKNESS E.F., PETERSON S.,  
RA LOFTUS B., RICHARDSON D., DODSON R., KHALAK H.G., GLODEK A.,  
RA MCKENNEY K., FITZGERALD L.M., LEE N., ADAMS M.D., HICKEY E.K.,  
RA BERG D.E., GOCAYNE J.D., UTTERBACK T.R., PETERSON J.D., KELLEY J.M.,  
RA COTTON M.D., WEIDMAN J.M., FUJII C., BOWMAN C., WATTHEY L., WALLIN E.,  
RA HAYES W.S., BORODOVSKY M., KARP P.D., SMITH H.O., FRASER C.M.,  
RA VENTER J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
pylori."  
RL NATURE 388:539-547(1997).  
CC -!- CATALYTIC ACTIVITY: ATP + SHIKIMATE = ADP + SHIKIMATE 3-PHOSPHATE.  
CC -!- PATHWAY: FIFTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE  
AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

EMBL; AE000536; G2313237; -  
DR TIGR; HP0157; -  
DR PROSITE; PS01128; SHIKIMATE\_KINASE; 1.  
DR PFAM; PF01202; SKI; 1.  
KW AROMATIC AMINO ACID BIOSYNTHESIS; TRANSFERASE; KINASE; ATP-BINDING.  
FT NP\_BIND 8 15 ATP (POTENTIAL).  
SQ SEQUENCE 162 AA; 18411 MW; 6A2AFFAA CRC32;

Query Match 33.3%; Score 63; DB 1; Length 162;  
Best Local Similarity 38.9%; Pred. No. 1.08e+00;  
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 129 FEKROALYEKNASFTIDA 146  
QY 7 FNKVTGLYKGEFVLIQS 24

RESULT 8  
ID DPOL\_HPBWD STANDARD; PRT: 788 AA.  
AC P17193;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.

OS DUCK HEPATITIS B VIRUS (WHITE SHANGHAI DUCK ISOLATE S31) (DHBV).  
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 90085807.

RA UCHIDA M., ESUMI M., SHIKATA T.;  
RT "Molecular cloning and sequence analysis of duck hepatitis B virus  
RT genomes of a new variant isolated from Shanghai ducks.";  
RL VIROLOGY 173:600-606(1989).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M325446; -  
DR PIR: B33746; JDVLD.  
DR PFAM: PF00078; rvt; 1.  
DR PFAM: PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM: PF00336; DNA\_pol\_viral\_C; 1.  
DR TRANSFERASE: DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.  
KW SEQUENCE 788 AA; 89264 MW; FA0BCD26 CRC32;  
SQ  
Query Match 33.3%; Score 63; DB 1; Length 788;  
Best Local Similarity 53.3%; Pred. No. 1.08e+00;  
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
Db 89 LGKLSGLYQMKGCSF 103  
QY 7 FNKVTGLYE-KGCEF 20  
RESULT 9  
ID DPOL\_HPBDD STANDARD; PRT; 836 AA.  
AC P03162;  
DT 21-JUL-1986 (REL. 01, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.  
OS DUCK HEPATITIS B VIRUS (DHBV).  
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 84138772.  
RA MANDART E., KAY A., GALBERT F.;  
RT "Nucleotide sequence of a cloned duck hepatitis B virus genome:  
RT comparison with woodchuck and human hepatitis B virus sequences.";  
RL J. VIROL. 49:782-792(1984).  
RN [2]  
RP SEQUENCE OF 794-836 FROM N.A.  
RX MEDLINE: 84216498.  
RA MOLNAR-KIMBER K.L., SUMMERS J.W., MASON W.S.;  
RT "Mapping of the cohesive overlap of duck hepatitis B virus DNA and of  
RT the site of initiation of reverse transcription.";  
RL J. VIROL. 51:181-191(1984).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: K01834; G325433; -  
DR PIR: A00710; JDVLD.  
DR PFAM: PF00078; rvt; 1.  
DR PFAM: PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM: PF00336; DNA\_pol\_viral\_C; 1.  
DR TRANSFERASE: DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.  
KW SEQUENCE 836 AA; 95275 MW; B2784CCE CRC32;  
SQ  
Query Match 33.3%; Score 63; DB 1; Length 836;  
Best Local Similarity 50.0%; Pred. No. 1.08e+00;  
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
Db 89 LGKLSGLYQMKGCTF 103  
QY 7 FNKVTGLYE-KGCEF 20

Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;  
Db 136 RHHLGKLSGLYQMKGCTF 153  
QY 4 RCTFNKVTGLYE-KGCEF 20  
RESULT 10  
ID DPOL\_HPBDD STANDARD; PRT; 787 AA.  
AC P30028;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.  
OS DUCK HEPATITIS B VIRUS (STRAIN CHINA) (DHBV).  
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 91045091.  
RA TONG S., MATTES F., TEUBNER K., BLUM H.E.;  
RT "Complete nucleotide sequence of a Chinese duck hepatitis B virus.";  
RL NUCLEIC ACIDS RES. 18:6139-6139(1990).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M21953; G325437; -  
DR PIR: S12841; JDVLD.  
DR PFAM: PF00078; rvt; 1.  
DR PFAM: PF00242; DNA\_pol\_viral\_N; 1.  
DR PFAM: PF00336; DNA\_pol\_viral\_C; 1.  
DR TRANSFERASE: DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.  
KW SEQUENCE 787 AA; 89214 MW; 4D88BA22 CRC32;  
SQ  
Query Match 32.8%; Score 62; DB 1; Length 787;  
Best Local Similarity 53.3%; Pred. No. 1.66e+00;  
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
Db 89 LGKLSGLYQMKGCTF 103  
QY 7 FNKVTGLYE-KGCEF 20  
RESULT 11  
ID DPOL\_HPBDD STANDARD; PRT; 788 AA.  
AC P17192;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
DE DNA POLYMERASE (EC 2.7.7.7).  
GN P.  
OS DUCK HEPATITIS B VIRUS (BROWN SHANGHAI DUCK ISOLATE S5) (DHBV).  
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 90085807.  
RA UCHIDA M., ESUMI M., SHIKATA T.;  
RT "Molecular cloning and sequence analysis of duck hepatitis B virus  
RT genomes of a new variant isolated from Shanghai ducks.";  
RL VIROLOGY 173:600-606(1989).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32590; G325450; -.
DR PIR; A33746; JDLVLD.
DR PFAM; PF00078; rvt; 1.
DR PFAM; PF00242; DNA_pol_viral_N; 1.
DR PFAM; PF00336; DNA_pol_viral_C; 1.
DR TRANSFERASE; DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION.
SQ SEQUENCE 788 AA; 89378 MW; EEB46243 CRC32;

Query Match 32.8%; Score 62; DB 1; Length 788;
Best Local Similarity 53.3%; Pred. No. 1.66e+00;
Matches 8; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 89 LGKLSGLYQMGCTF 103
Qy 7 FNKVTGLYE-KGCEF 20

RESULT 12
ID RSXB_BACSU STANDARD; PRT; 199 AA.
AC P17906;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SIGMA-B NEGATIVE EFFECTOR.
GN RSXB.
OS BACILLUS SUBTILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC BACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE; 91008924.
RX KALMAN S., DUNCAN M.L., THOMAS S.M., PRICE C.W.;
RT "Similar organization of the sigB and spoIIA operons encoding
RT alternate sigma factors of Bacillus subtilis RNA polymerase.";
RI J. BACTERIOL. 172:5575-5585(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA KASHARA Y., NAKAI S., LEE S., SADAIE Y., OGASAWARA N.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 1-179 FROM N.A.
RC STRAIN=168;
RX MEDLINE; 87109071.
RA DUNCAN M.L., KALMAN S.S., THOMAS S.M., PRICE C.W.;
RT "Gene encoding the 37,000-dalton minor sigma factor of Bacillus
RT subtilis RNA polymerase: isolation, nucleotide sequence, chromosomal
RT locus, and cryptic function.";
RL J. BACTERIOL. 169:771-778(1987).
RN [4]
RP FUNCTION.
RC STRAIN=PB2;
RX MEDLINE; 92276352.
RA BOYLAN S.A., RUTHERFORD A., THOMAS S.M., PRICE C.W.;
RT "Activation of Bacillus subtilis transcription factor sigma B by a
RT regulatory pathway responsive to stationary-phase signals.";
RL J. BACTERIOL. 174:3695-3706(1992).
RN [5]
RP FUNCTION.
RC STRAIN=PY22;
RX MEDLINE; 93224457.
RA BENSON A.K., HALDENWANG W.G.;
RT "Regulation of sigma B levels and activity in Bacillus subtilis";
RL J. BACTERIOL. 175:2347-2356(1993).
CC -!- FUNCTION: NEGATIVE EFFECTOR OF SIGMA-B OPERON EXPRESSION.
CC DEPHOSPHORYLATES RSBS.
CC -!- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34995; G143461; -.
DR EMBL; AB001488; D1020064; -.
DR EMBL; Z99106; E1182440; -.
DR PIR; C27762; C27762.
DR PIR; D36131; D36131.
DR SUTLISIT; BG10736; RSXB.
KW TRANSCRIPTION REGULATION; REPRESSOR; HYDROLASE.
FT DOMAIN 22 199 PP2C-LIKE.
SQ SEQUENCE 199 AA; 22144 MW; 148202EE CRC32;

Query Match 32.3%; Score 61; DB 1; Length 199;
Best Local Similarity 33.3%; Pred. No. 2.56e+00;
Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 138 YKTHATYKSGKFIHT 155
Qy 7 FNKVTGLYKGCFFVLOS 24

RESULT 13
ID UL07_HSV11 STANDARD; PRT; 296 AA.
AC P10191;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE PROTEIN UL7.
GN UL7.
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE;
OC ALPHAHERPESVIRINAE; SIMPLEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88274327.
RA MCGEOCH D.J., DALRYMPLE M.A., DAVIDSON A.J., DOLAN A., FRAME M.C.,
RA MCNAB D., PERRY L.J., SCOTT J.E., TAYLOR P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.";
RL J. GEN. VIROL. 69:1531-1574(1988).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL7,
CC EHV-1 55, VZV 53, EBV BBRF2, HCMV UL103 AND HVS-1 42.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10879; G221729; -.
DR EMBL; X14112; G59508; -.
DR EMBL; X14112; E312321; -.
DR PIR; G28133; WMBEX7.
SQ SEQUENCE 296 AA; 33059 MW; 480DFA53 CRC32;

Query Match 32.3%; Score 61; DB 1; Length 296;
Best Local Similarity 27.3%; Pred. No. 2.56e+00;
Matches 6; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Db 174 RTSFPERVRCFLRSCHWLVNT 195
Qy 3 KRCTFNKVTGLYKGCFFVLOS 24

RESULT 14
ID YAT1_METJA STANDARD; PRT; 313 AA.
AC Q58471;
```

```
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DI HYPOTHETICAL PROTEIN MJ1071.
GN MJ1071.
OS MATHANOCOCUS JANNASCHII.
OC ARCHAEA: EUYARCHAEOTA; MATHANOCOCCALES; MATHANOCOCACCACEAE;
OC MATHANOCOCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE: 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,
RA HERBEEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GEORGIAN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL SCIENCE 273:1058-1073(1995).
CC -!- SIMILARITY: TO THE C-TERMINAL OF M.JANNASCHII MJ0977.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M85164; G338035; -
CC PIR: A42093; A42093.
CC MIM: 600246; -
CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC PROSITE: PS00061; ETS_DOMAIN_3; 1.
CC PFAM: PF00178; Ets; 1.
CC HSSP: P14921; 2STW.
CC TRANSFAC: T02128; -
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
KW DNA_BIND 5 85
FT SEQUENCE 405 AA; 44674 MW; C66EC13E CRC32;
SQ
Query Match 32.3%; Score 61; DB 1; Length 405;
Best Local Similarity 35.0%; Pred. No. 2.56e+00;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
Db 216 YVKQSVNDKILGLEGEIQL 235
| : : | : | : | :
QY 1 YTKRCTFNKVTGLYKGEFVL 20
RESULT 15
ID SAPB_HUMAN STANDARD; PRT; 405 AA.
AC P28323;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE SERUM RESPONSE FACTOR ACCESSORY PROTEIN 1B (SAP-1B) (ETS-DOMAIN
DE PROTEIN ELK-4).
GN ELK4 OR SAP1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92154673.
RA DALTON S., TREISMAN R.;
RT "Characterization of SAP-1, a protein recruited by serum response
RT factor to the c-fos serum response element.";
RL CELL 68:597-612(1992).
CC -!- FUNCTION: FORMS A TERNARY COMPLEX WITH THE SERUM RESPONSE FACTOR
CC (SRF). REQUIRES DNA-BOUND SRF FOR TERNARY COMPLEX FORMATION AND
CC MAKES EXTENSIVE DNA CONTACTS TO THE 5' SIDE OF SRF, BUT DOES NOT
CC BIND DNA AUTONOMOUSLY.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: TWO FORMS OF SAP-1 ARE PRODUCED BY
CC -!- ALTERNATIVE SPLICING OF THE SAME GENE, THEY ONLY DIFFER IN THE
```

```
CC C-TERMINUS.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M85164; G338035; -
CC PIR: A42093; A42093.
CC MIM: 600246; -
CC PROSITE: PS00345; ETS_DOMAIN_1; 1.
CC PROSITE: PS00346; ETS_DOMAIN_2; 1.
CC PROSITE: PS00061; ETS_DOMAIN_3; 1.
CC PFAM: PF00178; Ets; 1.
CC HSSP: P14921; 2STW.
CC TRANSFAC: T02128; -
CC DNA-BINDING; NUCLEAR PROTEIN; ALTERNATIVE SPLICING.
KW DNA_BIND 5 85
FT SEQUENCE 405 AA; 44674 MW; C66EC13E CRC32;
SQ
Query Match 32.3%; Score 61; DB 1; Length 405;
Best Local Similarity 38.9%; Pred. No. 2.56e+00;
Matches 7; Conservative 3; Mismatches 7; Indels 1; Gaps 1;
Db 376 CPFKQIONLYTQVC-FLL 392
| : : | : | : | :
QY 5 CTENKVTGLYKGEFVL 22
Search completed: Mon Aug 23 13:05:09 1999
Job time : 10 secs.
```

**THIS PAGE BLANK (USPTO)**



\*\*\*\*\*

W P E R E H (TM)

\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Mon Aug 23 13:05:28 1999; Maspar time 12.65 Seconds  
Tabular output not generated. 103.568 Million cell updates/sec

Title: >US-09-049-696-43  
Description: (1-24) from US09049696.pcp  
Perfect Score: 189  
Sequence: 1 YTKRCTFNKVTGLYKGCFFVLQS 24

Scoring table: PAM 150  
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: spirembly  
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human  
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle  
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified  
13:sp-vertebrate 14:sp-virus

Statistics: Mean 28.880; Variance 44.497; scale 0.649

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	90	47.6	913	11	GOB-5 PROTEIN.	7.94e-05
2	74	39.2	785	14	POLYMERASE.	8.73e-02
3	73	38.6	272	2	DNAA (FRAGMENT).	1.32e-01
4	72	38.1	242	14	COAT PROTEIN.	2.00e-01
5	72	38.1	242	14	ORF5=25K.	2.00e-01
6	70	37.0	263	2	CONSERVED HYPOTHETICAL	4.51e-01
7	68	36.0	267	2	UNNAMED PROTEIN PRODUC	1.00e+00
8	68	36.0	267	2	LSG LOCUS HYPOTHETICAL	1.00e+00
9	68	36.0	267	2	ORF 6.	1.00e+00
10	68	36.0	357	2	GCPE PROTEIN.	1.00e+00
11	68	36.0	1000	6	TITIN (1000 AA) (FRAGM	1.00e+00
12	68	36.0	6875	6	TITIN (FRAGMENT).	1.00e+00
13	68	36.0	26926	4	TITIN, HEART ISOFORM N	1.00e+00
14	66	34.9	326	1	HOMOSERINE DEHYDROGENA	2.20e+00
15	65	34.4	261	2	GENE CLUSTER FOR 6-DEO	3.25e+00
16	65	34.4	592	2	SULFUR OXIDATION PROTE	3.25e+00
17	64	33.9	213	5	ZK484.6 PROTEIN.	4.77e+00
18	64	33.9	282	4	SERINE PROTEASE (TLSP)	4.77e+00
19	64	33.9	754	2	BETA-GLUCOSIDASE (EC 3	4.77e+00
20	63	33.3	65	14	ENVELOPE GLYCOPROTEIN	6.97e+00

21	63	33.3	94	14	070941	ENVELOPE GLYCOPROTEIN	6.97e+00
22	63	33.3	356	2	P72901	HYPOTHETICAL 41.6 KD P	6.97e+00
23	63	33.3	420	14	Q64844	47K PROTEIN.	6.97e+00
24	63	33.3	710	13	Q57520	GP130P3 (FRAGMENT).	6.97e+00
25	63	33.3	734	11	Q03930	PLATELET GLYCOPROTEIN	6.97e+00
26	63	33.3	786	14	Q03765	DNA-DIRECTED DNA POLYM	6.97e+00
27	63	33.3	836	14	Q71306	DNA POLYMERASE.	6.97e+00
28	63	33.3	836	14	Q66400	POLYMERASE.	6.97e+00
29	62	32.8	58	14	Q71600	ENVELOPE GLYCOPROTEIN	1.02e+01
30	62	32.8	245	2	Q52260	PLASMID PED208 SURFACE	1.02e+01
31	62	32.8	320	3	Q13842	HYPOTHETICAL 36.2 KD P	1.02e+01
32	62	32.8	436	10	Q04389	B-LIKE CYCLIN.	1.02e+01
33	62	32.8	788	14	Q72884	POLYMERASE.	1.02e+01
34	62	32.8	788	14	Q66403	P-PROTEIN.	1.02e+01
35	62	32.8	838	14	Q66402	COMPLETE GENOME.	1.02e+01
36	61	32.3	229	14	Q40081	GP120 (FRAGMENT).	1.47e+01
37	61	32.3	229	14	Q40087	GP120 (FRAGMENT).	1.47e+01
38	61	32.3	234	14	Q40084	GP120 (FRAGMENT).	1.47e+01
39	61	32.3	246	2	Q25244	HYPOTHETICAL 29.0 KD P	1.47e+01
40	61	32.3	319	1	Q28219	ACTIVATOR 1, REPLICATI	1.47e+01
41	61	32.3	340	5	P90627	CATHEPSIN B-LIKE PROTE	1.47e+01
42	61	32.3	432	2	Q87875	B-SUBUNIT OF BENZOYL-C	1.47e+01
43	61	32.3	781	10	Q80743	T13D8.9 PROTEIN.	1.47e+01
44	61	32.3	881	13	Q57519	GP130P1.	1.47e+01
45	61	32.3	950	2	Q51770	EXONUCLEASE SBCC (SBCC	1.47e+01

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	913	AA.
ID	O88826				
AC	O88826;				
DT	01-NOV-1998	(TREMBLREL. 08, CREATED)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	GOB-5 PROTEIN.				
GN	GOB-5				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;				
OC	SCIUROGNATHI; MURIDAE; MURINAE; MUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-INTESTINE;				
RA	KOMIYA T., TANIGAWA Y., HIROHASHI S.;				
RT	"Cloning of a gene, gob-5, which is expressed in intestinal goblet				
RT	cells in mice."				
RL	SUBMITTED (AUG-1998) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL: AB017156; D1034712;				
SQ	SEQUENCE 913 AA; 100071 MW; 4C9D2ECC CRC32;				

Query Match	47.6%	Score 90;	DB 11;	Length 913;
Best Local Similarity	50.0%;	Pred. No. 7.94e-05;		
Matches	9;	Conservative	4;	Mismatches 5; Indels 0; Gaps 0;
Db	210	KCVIDRVTGLYKDCNCFV 227		
QY	4	RCFKNVTGLYKGCFFV 21		

RESULT	2	PRELIMINARY;	PRT;	785	AA.
ID	Q67851				
AC	Q67851;				
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)			
DT	01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	POLYMERASE.				
OS	DUCK HEPATITIS B VIRUS (DBV).				
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; AVIHEPADNAVIRUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	SHI H., CULLEN J.M., NEWBOLD J.E.;				
RL	SUBMITTED (DEC-1992) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL: M95589; G325441;				

```

DR PFAM: PF00078; rvt; 1.
DR PFAM: PF00242; DNA_pol_viral_N; 1.
DR PFAM: PF00336; DNA_pol_viral_C; 1.
SQ SEQUENCE 785 AA; 89399 MW; B3CA4783 CRC32;

Query Match      39.2%; Score 74; DB 14; Length 785;
Best Local Similarity 55.6%; Pred. No. 8.73e-02;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Db 86 RHKLKSLGLYQMGCEF 103
      | : : : : : | : : : : |
QY 4 RCTFNKVTGLYE-KGCEF 20

RESULT 3
ID O54375 PRELIMINARY; PRT; 272 AA.
AC O54375;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE DNAA (FRAGMENT).
GN DNAA.
OS LACTOCOCCUS LACTIS (STREPTOCOCCUS LACTIS).
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC LACTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG1363;
RX MEDLINE; 98118563.
RA EL-KAROUI M., EHRlich D., GRUSS A.;
RT "Identification of the lactococcal exonuclease/recombinase and its
modulation by the putative Chi sequence.";
RL PROC. NATL. ACAD. SCI. U.S.A. 95:626-631(1998).
FR EMBL; U76424; G2909713; -.
FT NON_TER 1
SQ SEQUENCE 272 AA; 31014 MW; F9D10E6D CRC32;

Query Match      38.6%; Score 73; DB 2; Length 272;
Best Local Similarity 50.0%; Pred. No. 1.32e-01;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 54 FNTFNALYDKGSOIVLTS 71
      || : || : || : ||
QY 7 FNVTVGLYKGEFVLQS 24

RESULT 4
ID O37178 PRELIMINARY; PRT; 242 AA.
AC O37178;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE COAT PROTEIN.
OS BAMBOO MOSAIC VIRUS.
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; POTEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RA LIAO J.T., HSU Y.H.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF018156; G2407623; -.
DR PFAM: PF00286; virus_P-coat; 1.
KW COAT PROTEIN.
SQ SEQUENCE 242 AA; 25392 MW; B81B9454 CRC32;

Query Match      38.1%; Score 72; DB 14; Length 242;
Best Local Similarity 43.8%; Pred. No. 2.00e-01;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 165 KWCADFADFGLYDPTC 180
      | : | : | : | : |
QY 3 KRCTFNKVTGLYKGC 18

RESULT 5
ID O89760 PRELIMINARY; PRT; 242 AA.
AC O89760;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF5-25K.
OS BAMBOO MOSAIC VIRUS.
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; POTEXVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAMV-O;
RA HSU Y.;
RL SUBMITTED (DEC-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BAMV-O;
RX MEDLINE; 94358760.
RA LIN N.S., LIN B.Y., LO N.W., HU C.C., CHOW T.Y., HSU Y.H.;
RT "Nucleotide sequence of the genomic RNA of bamboo mosaic
potexvirus.";
RL J. GEN. VIROL. 75:2513-2518(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=GREEN BAMBOO;
RA LIN N.S., LIN B.Y., LO N.W., HU C.C., CHOW T.Y., HSU Y.;
RL J. GEN. VIROL. 75:0-0(0).
DR EMBL; D26017; G1228933; -.
DR EMBL; X57581; G509063; -.
DR PFAM: PF00286; virus_P-coat; 1.
KW COAT PROTEIN.
SQ SEQUENCE 242 AA; 25447 MW; 9A1B13F5 CRC32;

Query Match      38.1%; Score 72; DB 14; Length 242;
Best Local Similarity 43.8%; Pred. No. 2.00e-01;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 165 KWCADFADFGLYDPTC 180
      | : | : | : | : |
QY 3 KRCTFNKVTGLYKGC 18

RESULT 6
ID O51388 PRELIMINARY; PRT; 263 AA.
AC O51388;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN BB0427.
OS BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE; 98065943.
RA FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE O., KETCHUM K.A., DOBSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA VAN VUOT R., PALMER N., ADAMS M.D., COCAINE J.D., WEIDMAN J.,
RA UTTERBACK T., WATHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi".
RL NATURE 390:580-586(1997).
DR EMBL; AE001148; G2688342; -.
DR PROSITE; PS01296; UPF0011; 1.
DR TIGR; BB0427; -.
SQ SEQUENCE 263 AA; 29739 MW; F2166DC7 CRC32;

```

```

Query Match          37.0%; Score 70; DB 2; Length 263;
Best Local Similarity 43.8%; Pred. No. 4.51e-01;
Matches              7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 180 FKRIAEYKRGDAFVL 195
      :::: || :| |||
      7 FNRVTGLYKGCDFVL 22

RESULT 7
ID O05364 PRELIMINARY; PRT; 267 AA.
AC O05364;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE UNNAMED PROTEIN PRODUCT.
OS HAEMOPHILUS ACTINOMYCETEMCOMITANS.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC ACTINOBACILLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y4;
RX MEDLINE; 98084462.
RA YOSHIDA Y., NAKANO Y., YAMASHITA Y., KOGA T.;
RT "Identification of a genetic locus essential for serotype b-specific
RI antigen synthesis in Actinobacillus actinomycetemcomitans.";
RL INFECT. IMMUN. 66:107-114(1998).
DR EMBL; AB002668; D1020408; -.
DR PFAM; PF00535; Glycos_transf_2; 1.
SQ SEQUENCE 267 AA; 30266 MW; 87A0D870 CRC32;

Query Match          36.0%; Score 68; DB 2; Length 267;
Best Local Similarity 53.3%; Pred. No. 1.00e+00;
Matches              8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 150 KRCFPNHTVAYOKS 164
      ||| || :| :| :|
      3 KRCTFNKVTGLYKKG 17

RESULT 8
ID O05081 PRELIMINARY; PRT; 267 AA.
AC O05081;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE LSG LOCUS HYPOTHETICAL.
GN H11695.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95350630.
RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL SCIENCE 269:496-512(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA L TATUSOV R., MUSHEGIAN A.R., BORK P., BROWN N.P., HAYES W.S.,
RA BORODOVSKY M., RUDD K.E., KOONIN E.V.;
RA CURR. BIOL. 6:279-291(1996).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,
RA KERLAVAGE A.R., BULT C.J., TOMB J.F., DOUGHERTY B.A., MERRICK J.M.,
RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,
RA SCOTT J.D., SHIRLEY R., LIU L.I., GLODEK A., KELLEY J.M.,
RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,
RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,
RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEOGHAGEN N.S.M.,
RA GNEHM C.L., MCDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,
RA VENTER J.C.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RA WHITE O., CLAYTON R.A., KERLAVAGE A.R., FLEISCHMANN R.D.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U32842; G1574548; -.
DR PFAM; PF00535; Glycos_transf_2; 1.
SQ SEQUENCE 267 AA; 30770 MW; B66A686E CRC32;

Query Match          36.0%; Score 68; DB 2; Length 267;
Best Local Similarity 53.3%; Pred. No. 1.00e+00;
Matches              8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 150 KRCFPNHTVAYOKS 164
      ||| || :| :| :|
      3 KRCTFNKVTGLYKKG 17

RESULT 9
ID Q48215 PRELIMINARY; PRT; 267 AA.
AC Q48215;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF 6.
OS HAEMOPHILUS INFLUENZAE.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PASTEURELLACEAE;
OC HAEMOPHILUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MCLAUGHLIN R., ABU KWAIK Y., YOUNG R., SPINOLA S., APICELLA M.;
RL SUBMITTED (JUN-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; M94855; G148937; -.
DR PFAM; PF00535; Glycos_transf_2; 1.
SQ SEQUENCE 267 AA; 30681 MW; B80ABF16 CRC32;

Query Match          36.0%; Score 68; DB 2; Length 267;
Best Local Similarity 53.3%; Pred. No. 1.00e+00;
Matches              8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 150 KRCFPNHTVAYOKS 164
      ||| || :| :| :|
      3 KRCTFNKVTGLYKKG 17

RESULT 10
ID O67496 PRELIMINARY; PRT; 357 AA.
AC O67496;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GCPE PROTEIN.
GN GCPE.
OS AQUIFEX AEOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";

```

RL NATURE 392:353-358(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-VF5;  
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,  
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,  
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;  
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; AE000745; G2983917; -  
SQ SEQUENCE 357 AA; 39128 MW; 30231F62 CRC32;

Query Match 36.0%; Score 68; DB 2; Length 357;  
Best Local Similarity 56.3%; Pred. No. 1.00e+00;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 39 TLNQIKRLYEAGCEIV 54  
I: I: I I I I I I I  
Qy 6 TFNKVTGLYKGCERV 21

RESULT 11  
ID Q28737 PRELIMINARY; PRT: 1000 AA.  
AC Q28737;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE TITIN (1000 AA) (FRAGMENT).  
OS ORYCTOLAGUS CUNICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PSOAS MUSCLE;  
RX MEDLINE; 90238553.  
RA LABEIT S., BARLOW D.P., GAUTEL M., GIBSON T., HOLT J., HSIEH C.L.,  
RA FRANCKE U., LEONARD K., WARDALE J., WHITING A., TRINICK J.;  
RT "A regular pattern of two types of 100-residue motif in the sequence of titin.";  
RL NATURE 345:273-276(1990).  
DR EMBL; X17330; G930252; -  
DR PFAM; PF00041; fn3; 8.  
KW MUSCLE PROTEIN.  
FT NON\_TER 1  
FT NON\_TER 1000 1000  
SQ SEQUENCE 1000 AA; 110068 MW; 1F560957 CRC32;

Query Match 36.0%; Score 68; DB 6; Length 1000;  
Best Local Similarity 64.7%; Pred. No. 1.00e+00;  
Matches 11; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

Db 854 KCTY-KVTGLSE-GCEY 868  
I: I: I I I I I I I  
Qy 4 RCTFNKVTGLYKGCERV 20

RESULT 12  
ID Q28733 PRELIMINARY; PRT: 6875 AA.  
AC Q28733; Q28736;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE TITIN (FRAGMENT).  
OS ORYCTOLAGUS CUNICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CE12;  
RA LABEIT S.;  
RL SUBMITTED (FEB-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [2]  
RP SEQUENCE OF 1-6805 FROM N.A.  
RC STRAIN=CE12;

RX MEDLINE; 92258380.  
RA LABEIT S., GAUTEL M., LAKEY A., TRINICK J.;  
RT "Towards a molecular understanding of titin.";  
RL EMBL J. 11:1711-1716(1992).  
RN [3]  
RP SEQUENCE OF 4305-5320 FROM N.A.  
RC TISSUE=PSOAS MUSCLE;  
RX MEDLINE; 90238553.  
RA LABEIT S., BARLOW D.P., GAUTEL M., GIBSON T., HOLT J., HSIEH C.L.,  
RA FRANCKE U., LEONARD K., WARDALE J., WHITING A., TRINICK J.;  
RT "A regular pattern of two types of 100-residue motif in the sequence of titin.";  
RL NATURE 345:273-276(1990).  
DR EMBL; X64696; G1725; -  
DR EMBL; X17329; G930251; -  
DR PFAM; PF00041; fn3; 50.  
KW MYOSIN; MUSCLE PROTEIN.  
FT NON\_TER 1  
FT NON\_TER 6875 6875  
SQ SEQUENCE 6875 AA; 759127 MW; E5D3B61F CRC32;

Query Match 36.0%; Score 68; DB 6; Length 6875;  
Best Local Similarity 64.7%; Pred. No. 1.00e+00;  
Matches 11; Conservative 3; Mismatches 1; Indels 2; Gaps 2;

Db 2855 KCTY-KVTGLSE-GCEY 2859  
I: I: I I I I I I I  
Qy 4 RCTFNKVTGLYKGCERV 20

RESULT 13  
ID Q10466 PRELIMINARY; PRT: 26926 AA.  
AC Q10466;  
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
DE TITIN, HEART ISOFORM N2-B (EC 2.7.1.-) (CONNECTIN).  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;  
OC CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEART;  
RX MEDLINE; 96026330.  
RA LABEIT S., KOLMER B.;  
RT "Titins: giant proteins in charge of muscle ultrastructure and elasticity.";  
RL SCIENCE 270:293-296(1995).  
RN [2]  
RP SEQUENCE OF 22277-25376 FROM N.A.  
RX MEDLINE; 92258380.  
RA LABEIT S., GAUTEL M., LAKEY A., TRINICK J.;  
RT "Towards a molecular understanding of titin.";  
RL EMBL J. 11:1711-1716(1992).  
RN [3]  
RP SEQUENCE OF 1976-2014 FROM N.A.  
RA LABEIT S.;  
RL SUBMITTED (DEC-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [4]  
RP CHARACTERIZATION.

RX MEDLINE; 95331314.  
RA GAUTEL M., CASTIGLIONE-MORELLI M.A., PFUHL M., MOTTA A., PASTORE A.;  
RT "A calmodulin-binding sequence in the C-terminus of human cardiac titin kinase.";  
RL EUR. J. BIOCHEM. 230:752-759(1995).  
RN [1]  
RP FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.  
RN [2]  
RP TISSUE SPECIFICITY: MUSCLE-SPECIFIC.  
RN [3]  
RP ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE DISTRIBUTION.  
RN [4]  
RP DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY ONE TISSUE.



**This Page Blank (uspto)**